

DETERMINING SELECTIVE VOLUNTARY
MOTOR CONTROL OF THE LOWER EXTREMITY
IN CHILDREN WITH CEREBRAL PALSY

JULIA BALZER

A thesis submitted in partial fulfilment
of the requirements for the degree of
Doctor of Philosophy

QUEEN MARGARET UNIVERSITY

2018

Abstract

For physiotherapists working in neuro-paediatric gait-rehabilitation, improving motor control of the lower extremity is a major focus. Nevertheless, our understanding of selective voluntary motor control (SVMC) is in its infancy. This PhD project aimed to contribute to close this gap by investigating the nature of SVMC of the lower extremity in children with cerebral palsy (CP) and providing a psychometric robust yet sensitive measurement instrument for quantifying SVMC.

The first study investigated the influence of SVMC and other lower extremity and trunk motor impairments on gait capacity using multiple regression-analyses. Although SVMC was not kept within the final model, these study results revealed the importance of SVMC in relation to muscle strength, trunk control and gait capacity. The aim of the second study was to establish validity and reliability of the German version of the 'Selective Control Assessment of the Lower Extremity' (SCALE). Although the psychometric properties of the German SCALE were good, information about its responsiveness is lacking. Accordingly, a systematic review was carried out to identify a SVMC measurement instrument with the highest level of evidence for its psychometric properties and best clinical utility. As the findings showed the absence of appropriate, responsive SVMC measures, the aim of the last study was to modify the existing SCALE to make it more sensitive. Due to the positive findings in relation to the psychometric properties of the SCALE, its procedure was combined with a surface electromyography Similarity Index (SI). The first validity and reliability results of the SCALE-SI are promising and serve as benchmarks when applying the SCALE-SI in future clinical and scientific practice. However, to use the SCALE-SI as an outcome measure for detecting therapy-induced changes of SVMC in children with CP, its responsiveness needs to be evaluated in future studies.

Key Words: cerebral palsy, selective voluntary motor control, psychometric properties, lower extremity, gait rehabilitation

Acknowledgements

I would like to thank my supervising team, Dr. Marietta van der Linden, Professor Tom Mercer and Professor Hubertus van Hedel, for their guidance, feedback and motivation throughout my PhD (learning) process. Conducting and writing this PhD project was far more than an academic learning progress. Besides learning interesting scientific skills, I learned a lot about communication and interpersonal skills. Thereby it was uplifting for my brain and for my whole personality. This holistic engagement was only possible due to the special academic and personal support from my supervisors and their role-modelling. I am very grateful for these learning experiences which will serve me throughout my whole life.

Another very important support for fulfilment of this project has been the professional as well as motivational-emotional input from my research group colleges: *Corinne*: for being the brave PhD-Pioneer and my PhD-supporter from the beginning to the end; *Petra and Elena*: for working together with you 'Super-Girls' and especially for training my risible muscles; *Rob*: thank you very much for your on- and offline-EMG support right from the beginning; *Urs*: 'Professor of Matlab' thanks for planned and spontaneous Matlab support; *Jan, Jeff and Annina*: for the cooperative teamwork in organisation, recruitment and data analysis in wonderful "R"; *Judith, Nicole and Tabea*: for your special GCPs inputs; *Julia*: for all your precious practical-physio questions and inputs.

Furthermore, I would like to thank all in- and out-patients and parents for their participation. Regarding recruiting and data collection, I would like to especially thank all therapists of the "Vivendra-Stiftung", the "Schule für Körper- und Mehrfachbehinderte (SKB) Zürich", the "RGZ – Stiftung zugunsten cerebral Gelähmter" for their unbelievable great support, organisation and assistance in recruiting the participants. Only with their support it was possible to run all these projects. Thank you so much!!!

Last but not least, I would like to give my special thanks to the funder of these projects: the Swiss National Science Foundation, 'PhysioSwiss' (CH); 'Physiotherapy Science Foundation' (CH) and 'Mäxi-Foundation' (CH). We further acknowledge the Zurich Center for Neuroscience (ZNZ). Without their financial support my project-ideas would not have become reality. Thank you very much.

... and of course my family – my husband Nico, OUR children Kiki, Kelani and Keano and our dogs Romy and Capri – for your unconditional love (...no matter if I "make" it or not), their patience and support and all the thinking, as well as just being-present walks with the dogs. It was a wonderful and amazing process to see how our children have developed and grown since I have started with my PhD project and to have in parallel the possibility to work on this project.

... as well as my parents, who always believed in me and encouraged me to go for the things I am whole-hearted interested in.

... and lastly my brother for being connected with our heads and hearts.

Your vision will become clear only when you look inside your own heart
Who looks outside, dreams.
Who looks inside, awakes.
(C.G. Jung)

Table of Content

LIST OF TABLES.....	IX
LIST OF FIGURES.....	X
LIST OF APPENDICES.....	XI
ABBREVIATIONS.....	XII
PUBLICATIONS ARISING FROM THIS THESIS.....	XIV
PRESENTATIONS ARISING FROM THIS THESIS	XV
CHAPTER 1: INTRODUCTION	1
1.1 Purpose of chapter.....	1
1.2 “Will my child learn to walk normally?” – My ethical and professional dilemma	1
1.3 Motor control problems in children with cerebral palsy	2
1.4 Importance of physiotherapy in training motor control	3
1.4.1 Outcome measurement within neuro-paediatric physiotherapy	5
1.5 Structure of thesis	6
CHAPTER 2: CHAPTER: BACKGROUND.....	8
2.1 Purpose of chapter.....	8
2.2 Cerebral Palsy and selective voluntary motor control (SVMC)	8
2.2.1 Epidemiology of cerebral palsy.....	8
2.2.2 Aetiology of cerebral palsy	9
2.2.3 Classification of cerebral palsy	10
2.2.4 Underlying pathophysiology of impaired SVMC in children with CP....	12
2.2.4.1 Maturation, functioning and impairment of the corticospinal tract ...	13
2.2.5 Clinical symptoms of corticospinal tract lesions	21
2.2.6 Altered processes, structures and networks in relation to impaired SVMC.....	24
2.2.7 Impaired SVMC in children with CP along the ICF continuum	26
2.2.7.1 SVMC impairments on body functions and structures level	27
2.2.7.2 Possible SVMC caused limitations on the activity level	28
2.2.7.3 Possible SVMC caused disabilities on participation level	29
2.2.8 Trainability of SVMC.....	32
2.2.9 Neuronal structures and mechanisms involved in improving impaired SVMC	36

2.2.9.1	Concept of neuroplasticity	36
2.2.9.2	Neurophysiology of constraint-induced therapy and mirror therapy	37
2.3	Outcome measures.....	39
2.3.1	Psychometric properties	41
2.3.2	Measuring the quality of studies presenting psychometric data	42
2.3.3	Reliability.....	44
2.3.4	Validity	46
2.3.5	Responsiveness.....	48
2.3.6	Outcome measures for measuring SVMC	49
2.3.6.1	Measures of SVMC in children with CP.....	50
2.3.6.2	Measures of SVMC in other populations with a UMN lesion.....	51
2.4	Rationale for the overall aim(s) of the thesis.....	54
CHAPTER 3:	STUDY 1	58
3.1	Purpose of chapter.....	58
3.2	Background.....	58
3.3	Method.....	60
3.3.1	Participants	60
3.3.2	Measurements	60
3.3.3	Lower extremity assessments	60
3.3.4	Trunk control assessment	61
3.3.5	Gait capacity assessment.....	62
3.3.6	Statistical Analysis.....	62
3.4	Results.....	63
3.4.1	Correlation analysis.....	63
3.4.2	Simple and multiple linear regression analysis	65
3.5	Discussion	67
3.5.1	Prediction of gait capacity.....	67
3.5.2	Relationship between lower extremity motor functioning and trunk control	68
3.5.3	Limitations and Methodological Considerations.....	69
3.5.4	Clinical implications and future work.....	71
3.6	Conclusion	72

CHAPTER 4: STUDY 2	73
4.1 Purpose of chapter	73
4.2 Background.....	73
4.3 Method.....	74
4.3.1 Participants	74
4.3.2 Measurements	74
4.3.3 Statistical Analysis.....	76
4.4 Results.....	77
4.4.1 Discriminative validity	78
4.4.2 Correlations.....	81
4.4.3 Reliability.....	82
4.5 Discussion	82
4.5.1 Validity	82
4.5.2 Reliability.....	84
4.5.3 Limitations and Methodological Considerations.....	84
4.5.4 Clinical implications and future work.....	84
4.6 Conclusion	85
CHAPTER 5: STUDY 3	86
5.1 Purpose of chapter.....	86
5.2 Background.....	86
5.3 Method.....	87
5.3.1 Search Strategy	87
5.3.2 Study Selection	87
5.3.3 Quality Evaluation	88
5.4 Results.....	89
5.4.1 Description of the included studies	89
5.4.2 Hypotheses testing.....	91
5.4.3 Content and Criterion Validity	93
5.4.4 Reliability.....	93
5.5 Discussion	100
5.5.1 SVMC Assessment Tools.....	100
5.5.2 SVMC laboratory-based Measures.....	101
5.5.3 Limitations and methodological considerations.....	101
5.5.4 Clinical implications and future work.....	102
5.6 Conclusion	103

CHAPTER 6: STUDY 4	104
6.1 Purpose of chapter	104
6.2 Background.....	104
6.3 Method.....	106
6.3.1 Participants	106
6.3.2 Measurements	107
6.3.3 Data recording and processing.....	109
6.3.4 Statistical Analysis.....	111
6.4 Results.....	112
6.4.1 Validity	114
6.4.2 Reliability.....	117
6.5 Discussion	119
6.5.1 Validity	120
6.5.2 Reliability.....	124
6.5.3 Limitations and methodological considerations.....	125
6.5.4 Clinical Implications and future work	126
6.6 Conclusion	127
CHAPTER 7: FINAL DISCUSSION	128
7.1 Purpose of chapter.....	128
7.2 Summary of main findings.....	128
7.2.1 Main findings study 1.....	128
7.2.2 Main findings study 2.....	129
7.2.3 Main findings study 3.....	129
7.2.4 Main findings study 4.....	130
7.3 Considerations when testing psychometric properties in SVMC measures	131
7.3.1 Considerations when testing validity of SVMC measure.....	131
7.3.1.1 Considerations for criterion validity.....	132
7.3.1.2 Considerations for hypotheses testing: convergent and discriminative validity	135
7.3.2 Considerations when testing reliability of a SVMC measure	146
7.3.3 Considerations concerning clinical utility of SVMC measures.....	148

7.3.4	Considerations when testing responsiveness in SVMC measures....	149
7.3.4.1	Finding an adequate intervention for responsiveness testing	150
7.3.4.2	Outlook study design for responsiveness testing of the SCALE-SI	152
7.4	Common methodological considerations, limitations and future work	157
7.4.1	Multicollinearity.....	157
7.4.2	Generalizability is limited to children with GMFCS levels I-III.....	159
7.4.3	Estimation of clinical meaningful threshold	159
7.4.4	Future work	160
7.5	Conclusion	161
REFERENCES	163
APPENDICES	186

List of Tables

Table 1:	Interventions aiming to improve SVMC of the lower extremity in children with CP	34
Table 2:	Domains and definitions of measurement properties.....	42
Table 3:	Quality criteria for measurement properties.....	43
Table 4:	Study 1 - Participants' clinical and functional characteristics	63
Table 5:	Study 1 - Spearman's correlation coefficients (ρ) between lower extremity impairments, trunk control and gait capacity	65
Table 6:	Study 1 - Simple and multiple linear regression analysis for predicting gait capacity.....	66
Table 7:	Study 2 - Participants' characteristics.....	78
Table 8:	Study 2 - Intra- and Inter-rater reliability of the SCALE	82
Table 9:	Study 3 - General characteristics, psychometric properties and clinical utility SVMC measures	95
Table 10:	Study 3 - Summary study details, COMSIN-, quality and evidence-rating.....	96
Table 11:	Study 4 – Participants' characteristics.....	113
Table 12:	Study 4 – Norm and control group characteristics	114
Table 13:	Study 4 - Correlations SCALE scores vs SCALE SI	115
Table 14:	Study 4 - Differences neurologically intact children vs. children with CP	116
Table 15:	Study 4 - Relative and absolute reliability SCALE-SI	117

List of Figures

Figure 1:	Structure of thesis:	7
Figure 2:	Surveillance of Cerebral Palsy in Europe (2013): Classification tree for sub-types of CP	11
Figure 3:	Periventricular Leukomalacia and somatotopic organisation of the motor cortex, showing the close proximity of the lesion, CST and motor cortex.....	18
Figure 4:	Possible symptoms due to an UMN lesion: showing its complexity and long-term vicious circle effect in children with CP.	22
Figure 5:	International Classification of Functioning, Disability and Health (ICF) continuum:	27
Figure 6:	Clinical consequences of impaired SVMC presented at the ICF levels:	31
Figure 7:	Study 1 - Scatterplots:.....	64
Figure 8:	Study 2 - Discriminative validity for children with spastic CP:	80
Figure 9:	Study 2 – Spearman’s rank correlation coefficient (ρ) of total SCALE scores and common clinical assessments for children with spastic CP:.....	81
Figure 10:	Study 3 - Flow chart systematic review	90
Figure 11:	Study 4 - SCALE-SI testing setup:	108
Figure 12:	Study 4 - Correlation total SCALE score and GMFCS levels.....	115
Figure 13:	Study 4 – Hypotheses proximal-distal-concordance:.....	116
Figure 14:	Study 4 – SCALE-SI grouped by leg involvement and diagnosis:...	117
Figure 15:	Study 4 - Bland-Altman Plots: Test retest reliability SCALE-SI – less affected side	118
Figure 16:	Study 4 - Bland-Altman Plots: Test retest reliability SCALE-SI - more affected side:	119
Figure 17:	Study 4 - SCALE-SI:	121
Figure 18:	Study 4 - SCALE vs. SCALE-SI	122
Figure 19:	SCALE-SI vs. age	123
Figure 20:	Concurrent validity: SCALE vs. SCALE-SI on joint level.....	135
Figure 21:	Proximal-distal concordance:	139
Figure 22:	Correlations SCALE, SCALE-SI vs MAS:.....	141
Figure 23:	Correlations between SCALE / SCALE-SI and GMFCS:	145

List of Appendices

Appendix 1: COSMIN checklist.....	187
Appendix 2: Search-Construct SVMC measurement-tool list	210
Appendix 3: Overview of SVMC measurement-tools.....	212
Appendix 4: Study - German Version SCALE	214
Appendix 5: Study 2 - English SCALE version (original)	220
Appendix 6: Study 3 - Search terms for all databases.....	224
Appendix 7: Study 3 - established specified COSMIN rating rules	230
Appendix 8: Study 3 - Quality Criteria for Measurement Properties	232
Appendix 9: Study 3 - Evidence Level of Measurement Properties	233
Appendix 10: Publication study 1	234
Appendix 11: Publication study 2.....	242
Appendix 12: Publication study 2.....	249

Abbreviations

β	Beta (standardized regression coefficient)
CPGs	Central Pattern Generators
COSMIN	COnsensus-based Standards for the selection of health Measurement Instruments
CP	Cerebral Palsy
CST	Corticospinal Tract
d	differences between the ranks
FES	Functional Electrical Stimulation
GPS	Gait Profile Score
GM	m. Gastrocnemius Medialis
ICF	International Classification of Functioning, Disability and Health
ICC	Intraclass Correlation Coefficient
IQR	Interquartile Range
k	number of raters / measurements
m	mean
MAS	Modified Ashworth Scale
Mdn	Median
MDC	Minimal Detectable Change
MEP	Motor Evoked Potential
MMT	Manual Muscle Testing
MS _c	Mean Square for columns
MS _E	Mean Square for errors
MS _R	Mean Square for rows
mTUG	Modified Timed Up and Go Test
n	number of observations
PL	m. Peroneus Longus
PRV	Prototyp Response Vector
PVL	Periventricular Leukomalacia
R	Regression
<i>r</i>	Effect size
RF	m. Rectus Femoris
ROC	Receiver Operating Characteristic
ρ	Spearman's correlation coefficients (rho)

RST	Rubospinaltract
RV	Response Vector
SCALE	Selective Control Assessment of the Lower Extremity
sEMG	Surface Electromyography
SCPE	Surveillance of Cerebral Palsy in Europe
SEM	Standard Error of Measurement
SD	Standard Deviation
SDC	Smallest Detectable Change
SI	Similarity Index
SMC	Selective Motor Control
ST	m. Semitendinosus
SVMC	Selective Voluntary Motor Control
TA	m. Tibialis Anterior
TCMS	Trunk Control Measurement Scale
TMS	Transcranial Magnetic Stimulation
<i>U</i>	Mann-Whitney-U-test
UMN	Upper Motor Neuron
VIF	Variance Inflation Factor
VRI	Voluntary Response Index
Z	Z value (Wilcoxon Signed Ranks Test)

Publications arising from this thesis

Balzer J, Marsico P, Mitteregger E, van der Linden ML, Mercer TH, van Hedel HJ. 2017. Influence of trunk control and lower extremity impairments on gait capacity in children with cerebral palsy. *Disabil Rehabil.* 24:1-7.

Balzer J, van der Linden ML, Mercer TH, van Hedel HJ. 2017. Selective voluntary motor control measures of the lower extremity in children with upper motor neuron lesions: A systematic Review. *Dev Med Child Neurol.* 59(7):699-705.

Marsico P, Mitteregger E, Balzer J, van Hedel HJ. 2017. The Trunk Control Measurement Scale: reliability and discriminative validity in children and young people with neuromotor disorders. *Dev Med Child Neurol.* 59(7):706-712.

Balzer J, Marsico P, Mitteregger E, van der Linden ML, Mercer TH, van Hedel HJ. 2016. Construct validity and reliability of the Selective Control Assessment of the Lower Extremity in children with cerebral palsy. *Dev Med Child Neurol.* 58(2):167-72.

Marsico P, Frontzek-Weps V, Balzer J, van Hedel HJ. 2016. Hypertonia Assessment Tool: Reliability and Validity in Children with Neuromotor Disorders. *J Child Neurol.* 32(1):132-138.

Mitteregger E, Marsico P, Balzer J, van Hedel HJ. 2015. Translation and construct validity of the Trunk Control Measurement Scale in children and youths with brain lesions. *Res Dev Disabil.* 45-46:343-52.

Presentations arising from this thesis

“Influence of trunk control and lower extremity impairments on gait capacity in children with cerebral palsy” European Academy of Childhood Disability, Amsterdam (NL), 20th May 2017.

“Measuring selective voluntary motor control of the lower extremity in children with upper motor neuron lesions – A COSMIN review” International Conference on Cerebral Palsy and other Childhood Onset Disabilities, Stockholm (S), 4th July 2016.

“Validity and reliability of the German version of the "Selective Control Assessment of the Lower Extremity" (SCALE) in children with cerebral palsy” European Paediatric Neurology Society, Wien (AU); 28th May 2015.

Chapter 1: Introduction

1.1 Purpose of chapter

The aim of this chapter is to describe the practical as well as scientific context in which this thesis has been arisen.

1.2 “Will my child learn to walk normally?” – My ethical and professional dilemma

“Will my child learn to walk normally due to your therapy?” or “Can you teach him to walk normally?” These kinds of questions are common questions parents asked me during the rehabilitation stay of their child at our rehabilitation centre in Affoltern am Albis in Switzerland. Although our centre is specialized in neuro-paediatric rehabilitation and I am a physiotherapist specialized in rehabilitating children’s gait, the answer to such questions cannot easily be given. Of course, as a therapist, I wish to answer these questions in an honest and clear way. Mainly to save parents and their child from having unrealistic therapy aims and thereby wasting their time. Exactly this ethical and professional dilemma became the starting point of my thesis. Firstly, this dilemma led me to search for practical as well as scientific facts which would guide me to an answer. During my practical work, I discovered that improvements of selective voluntary movement control (SVMC) of the lower extremity, as one fundamental prerequisite for a physiological gait pattern, could be achieved in some patients but not in others, or only to a minimal degree. Discussions with my colleagues supported this clinical observation. Nevertheless, the standard problem we experienced was how to measure changes in SVMC objectively and accurately. A robust, e.g. valid, reliable and responsive, as well as sensitive outcome measure for SVMC was necessary.

Scientific literature revealed similar positive intervention trends regarding the trainability of SVMC, but also the lack of robust outcome measure. As there is no gold standard for measuring SVMC, a variety of measures, have been applied in these intervention studies, often with unknown psychometric qualities.

On the assumption that improved SVMC would contribute to the 'normalisation' of the child’s gait pattern, the aim of my PhD was born: to find a validated instrument to measure therapy-induced changes of SVMC of the lower extremity in children with

CP. The following paragraphs contain a detailed introduction to the topic of SVMC and the structure of this PhD, thereby explaining the various difficulties associated with predicting the probability of improving gait function and quality through physiotherapy.

1.3 Motor control problems in children with cerebral palsy

“Cerebral Palsy (CP) describes a group of permanent disorders of the development of movement and posture, causing activity limitations that are attributed to a non-progressive disturbance that occurred in the developing foetal or infant brain. The motor disorders of cerebral palsy are often accompanied by disturbance of sensation, perception, cognition, communication and behaviour, by epilepsy and by secondary musculoskeletal problems.” (Baxter et al. 2008). With an incidence of 2 to 2.5 per 1000 children born alive (SCPE 2000), CP is the most common childhood motor disorder (Baxter et al. 2008). There are various neuropathological causes for CP (e.g. perinatal stroke, hypoxia-Ischaemia, encephalopathy, cerebral malformation), which can affect different areas of the brain (Dan et al. 2014). Depending on the neuropathology and injured neuroanatomical structure, motor control is disturbed to various extents, and different motor function impairments occur. From a motor control perspective, CP can be seen as a group of movement disorders which is primarily caused by impaired cortical control, and further restricts the central as well as peripheral sensori-motor system (Rosenbaum 2014). Therefore, the motor development of children with CP is disturbed and delayed. Primarily, the lesion in the white matter disrupts the cortico-cortical and cortico-spinal pathways and thereby impedes the child’s motor control (Forssberg 2014). Due to injured cortical structure, the cortical motor output cannot reach the motor neurons in the spinal cords. This leads to a further alteration of the spinal interneuron networks, causing disturbed motor output as well as pathological changes within the muscle (Rose and McGill 2005; Miller 2007). Consequently, the child has problems with learning and executing movements in a coordinated, physiological way. Developmentally, this lack of coordination leads to further alterations in the maturation of the corticospinal tract (Forssberg 2014; Peacock 2009). While in neurologically intact children the bilateral corticospinal projections from the motor cortex to the extremity change into a contralateral projection within the first year (of living), this maturation process is disturbed in children with CP (Eyre et al. 2001; Carr et al. 1993; Staudt 2007). Clinically, these alterations in cortical control make it difficult for the child to selectively

activate one muscle or muscle group. Co-movements (simultaneous movements of adjacent joints of the ipsilateral side) and/or mirror-movements (simultaneous movements of the contralateral joint), as well as mass-patterns (e.g. extension-pattern: simultaneous hip, knee and foot extension) will occur and restrict the child's goal-directed movement (Cahill-Rowley and Rose 2014; Fowler 2010). For instance, selectively extending the knee and simultaneously dorsiflexing the ankle – as required for initial contact during gait – can be disturbed by a mass extension pattern, which leads to the simultaneous extension of the knee and the ankle, resulting in an “equinus” gait pattern (Gage et al. 2009). Furthermore, this impaired motor control is connected with the appearance of other upper motor neuron (UMN) lesion signs such as hypertonia, hyper-reflexia, muscle weakness, sensory disturbance, and sensory deficits (Cahill-Rowley and Rose 2014; Carr 2014; Miller 2007; Sanger et al. 2006). All these UMN lesion signs can reinforce each other, hampering motor control and development of the child even more and furthermore leading to the occurrence of secondary deformities (e.g. contractures) (Gage et al. 2009; Peacock 2009; Miller 2007). This interconnection between SVMC and other UMN signs (spasticity, muscle weakness, sensory deficits, balance problems) in children with CP makes it challenging for a therapist to measure and treat SVMC separately from other UMN impairments (Cahill-Rowley and Rose 2014; Dobson 2010). As a clear distinction of SVMC from other UMN signs is fundamental for the valid measurement of SVMC, these issues are addressed in more detail in section 2.2 on the neurophysiological background of SVMC in CP.

1.4 Importance of physiotherapy in training motor control

In neuro-rehabilitation, one major physiotherapeutic goal is to restore the patient's motor control as best as possible to maximize participation in society and quality of life (Mayston 2014; Miller 2007; Gage et al. 2009). Especially in children with CP, medical, therapeutic and financial health-care efforts are high in order to reach the best possible outcome. Due to the congenital nature of the brain lesion, patients can benefit from natural development and brain plasticity during rehabilitation which can enhance their motor recovery (Mayston 2001). On the other hand, as a diagnosis of CP is commonly associated with a normal or near normal lifespan, these patients are vulnerable to develop secondary deformities and chronic pain. Due to a combination of altered developmental growth factors (e.g. endocrine factors) and impaired motor control, biomechanical loading of muscles and bones is altered during movement and

posture right from the beginning of life. Over time, this non-physiological biomechanical stress causes muscle contractures and/or bony deformities to develop (Miller 2007; Gage et al. 2009). These contractures can worsen over time and thereby limit the child's movement ability even more. This vicious cycle of impaired motor control and growth can cause pain and limit the patient's quality of life. In many patients, (multiple) surgeries may be required during childhood/puberty (and adulthood) in order to realign muscles and bones physiologically and to relieve pain. Considering these two aspects, intensive rehabilitation and gait-therapy in children with CP are medically and economically essential.

Since the work of Berta and Karel Bobath and other pioneers within neuro-physiotherapy, several therapy concepts were developed and aimed to (re-)train physiological movements in patients with central nervous system lesions (Levin and Panturin 2011; Barber 2008; Mayston 2014). At its earliest inception, the Bobath concept focused on regaining normal movements through re-education. Its key techniques, namely therapeutic handling, facilitation, and activation of key points of control, aimed to improve patients' motor control while using the different stages of normal motor development as guidance. Although this approach arose from their best intention (e.g. reducing the risks of secondary deformities) and the available knowledge at that time, new information about motor learning and neuroplasticity evolved and led to the new paradigm "Activity, activity, activity" (Damiano 2006). This paradigm promoted a shift from traditional and strict Bobath approaches to a proactive approach of promoting activity through more intense active training protocols, lifestyle modifications and mobility-enhancing devices (Mantovani and Scrutton 2014). Research showed that allowing children with CP to be more active and to move more while accepting compensatory/non-physiological movements resulted in improved physical and mental health (Mayston 2014). Although, nowadays the new "Activity" paradigm (has) arrived within neuro-paediatric physiotherapy, therapists are still challenged by the balancing act between allowing the patient to be more active, teaching physiological movement patterns, and preventing secondary impairments. As these secondary impairments (e.g. contractures) are known to develop as consequences from moving within non-physiological movement patterns, teaching a patient physiological movement patterns is still of importance (Vos et al. 2016; Levin 2010; Voorman et al. 2007). If a patient, for instance, is not able to selectively activate the quadriceps to achieve full knee extension in gait, a knee-flexion contracture might develop, which hampers the ability to stand and walk even more. In the long term,

walking with a flexion contracture will lead to altered muscle-bone alignment and muscle insufficiency and may cause chronic pain. Consequently, this vicious cycle, which originated from impaired motor control and development, frequently results in reduced motor activity (Miller 2007; Gage et al. 2009).

Despite these extensive clinical consequences of impaired SVMC and the traditional therapeutic interest in its training, scientific investigation of the efficacy of physiotherapy interventions to improve selective muscle activation has only started recently. The first of these studies have shown that SVMC of the lower extremity has improved as a result of Functional Electrical Stimulation (FES) and virtual reality enhanced (robotic) training as well as due to medical spasticity reduction (e.g. Botulinum-toxin, Selective Dorsal Rhizotomy) in combination with conventional physiotherapy (Rios et al. 2013; Sukal-Moulton, Clancy, L.-Q. Zhang, et al. 2014; Cioi et al. 2011; Approach 2011; Prosser et al. 2012; Bandholm et al. 2009; Damiano et al. 2012). Furthermore, research into the mechanisms of neuroplasticity supports the trainability of SVMC in neurologically impaired patients (Forssberg 2014; Nudo 2013; Caeyenberghs et al. 2010; Everaert et al. 2010). The lack of a psychometrically sound gold standard measure or agreed core-set of outcome measures for SVMC may be an important reason for this gap in the physiotherapy literature. It is our professional duty to close this gap of evidence, in order to ensure our clients and their family the best possible treatment. Only by the means of valid, reliable and responsive measurement tools, we will be able to investigate the efficacy of treatment approaches aiming to improve motor control. Accordingly, the next paragraph will explain the different types of measures and how their quality can be evaluated.

1.4.1 Outcome measurement within neuro-paediatric physiotherapy

Measuring and evaluating intervention outcome(s) is the cornerstone of good clinical and scientific practice. Although this statement sounds clear and logical, implementation into clinical and research practice is often not that straight forward. Often the professional is confronted with the challenge to find “the right tool for the job” (Rosenbaum 2014). If it is the interest to measure the effectiveness of an intervention, the purpose of the measure is to detect a change in the outcome of primary importance on the body functions and structures, activity or participation domains according to the International Classification of Functioning Disability and Health (ICF) continuum (Wright and Majnemer 2014). Furthermore, the level of measurement (e.g. ordinal, interval) needs to be equivalent to the expected change

of the intervention within the population of interest (Portney and Watkins 2000). Additionally, the measurement tool needs to have good psychometric properties (De Vet et al. 2011). Psychometric properties describe the measurement properties such as the internal consistency, reproducibility or reliability, validity and responsiveness of the outcome measure (De Vet et al. 2011). To help researchers and clinicians with the evaluation of psychometric properties, the COSMIN initiative (COnsensus-based Standards for the selection of health Measurement INstruments) provides a consensus about the terminology and definitions of these measurement properties and guidelines for their evaluation (Mokkink and Terwee 2010). Only by precisely knowing the benefits and limitations of the outcome measure used, appropriate evaluation of the intervention will be possible. A discussion of the evaluation of psychometric properties in accordance with the COSMIN guidelines will be presented in greater detail in chapter 2, as these are fundamental for this thesis.

Although within the last years the number of measurement tools within the field of CP has increased in all domains of the ICF, there are still some outcome measures missing or their psychometric properties have not sufficiently been evaluated yet (Wright and Majnemer 2014; Mayston 2001). This seemed to be the case for a tool which measures therapeutically induced changes of SVMC of the lower extremity in children with CP. Therefore, the overarching aim of this thesis is to identify or to develop a psychometrically sound outcome measure for this purpose.

1.5 Structure of thesis

The thesis is structured as followed (Figure 1). The background chapter is divided into two main sections: Section one presents a detailed description of the population of interest with a focus on the pathophysiology of impaired SVMC and its functional and therapeutically consequences. Section two will summarise common problems when measuring SVMC, furthermore it will define relevant psychometric property quality criteria for outcome measures.

Chapter 3 presents the methods and results of the first study. This study investigated the influence and the relevance of SVMC and other UMN motor impairments on gait capacity in children with CP. A description of the second study, which dealt with the translations and validation process of a recently developed clinical tool for the assessment of SVMC ("Selective Control Assessment of the Lower Extremity" (SCALE)) of the lower extremity in children with CP into German, is described in

chapter 4. The methods and results of a systematic review of the psychometric properties of measurement tools for SVMC of the lower extremity in children with CP are presented in chapter 5. In chapter 6, validation and reliability of a newly developed outcome measure for SVMC (SCALE-Similarity Index (SI)) is described.

The final discussion chapter brings together the conclusions from each study and presents their implications in the context of neuropediatric gait-rehabilitation and research. Moreover, it relates to the start-up question of this PhD-project: “Will my child learn to walk normally due to your therapy?”

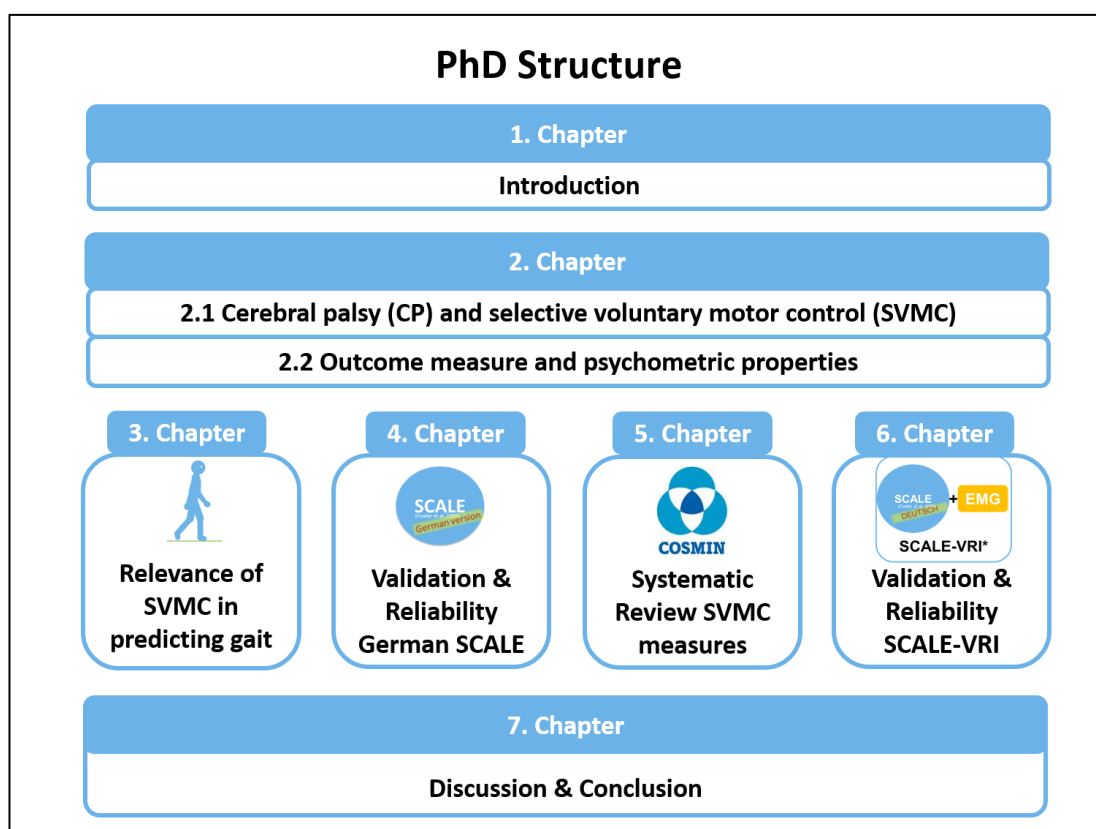


Figure 1: Structure of thesis:

Following a description of the theoretical background in relation to CP, SVMC and psychometric properties of outcome measures, the four studies, which have been carried out within the scope of this thesis, will be presented. The thesis ends with an overall discussion and conclusion.

Chapter 2: Chapter: Background

2.1 Purpose of chapter

The purpose of this chapter is to summarise the existing evidence as well as to identify gaps of knowledge in relation to the measurement of SVMC in children with CP. Findings in literature in relation to the focus of this thesis will be discussed to support the rationale behind this thesis. Consequently, the research questions will be presented.

2.2 Cerebral Palsy and selective voluntary motor control (SVMC)

In the “International Statistical Classification of Diseases and Related Health Problems” (ICD) diagnosis code, cerebral palsy is defined as

“(G80) A group of disorders affecting the development of movement and posture, often accompanied by disturbances of sensation, perception, cognition, and behaviour. [...] A heterogeneous group of non-progressive motor disorders caused by chronic brain injuries that originate in the prenatal period, perinatal period, or first few years of life. The four major subtypes are spastic, athetoid, ataxic, and mixed cerebral palsy, with spastic forms being the most common. The motor disorder may range from difficulties with fine motor control to severe spasticity (see muscle spasticity) in all limbs. [...] Pathologically, this condition may be associated with leukomalacia periventricular. [...]”(CD-10-CM 2018).

Although CP is a clinical diagnose, its diagnosis differs from other neurodevelopmental conditions, as accurate/sensitive diagnostic tests are missing (Rosenbaum 2014). Therefore, CP is better understood as an umbrella term for a heterogeneous group of developmental movement and posture impairments affecting the child's independence. The following paragraph's aim is to explain the epidemiology and aetiology of this group of motor disorders with a specific focus on the underlying pathophysiology and functional consequences of their impaired motor control.

2.2.1 Epidemiology of cerebral palsy

As the confirmation of the clinical diagnose of CP relies more on the systematic clinical observation of the child's motor development and less on pathognomic findings or laboratory test, collecting epidemiological data is challenging (Rosenbaum 2014; Eunson 2016). Furthermore, as CP is associated with an increased risk of infant mortality especially in low- and middle-income countries, information about its frequency and risk factors in those countries might be incomplete. Despite these

limitations, the following findings regarding the prevalence of CP can be summarised from the literature (Oskoui et al. 2013):

- With an overall birth prevalence of around 2 per 1,000 live births, CP is the most common childhood disability of high-income countries (Eunson 2016).
- CP has a higher prevalence in more deprived socio-economic populations (Rosenbaum 2014).
- Although maternal and neonatal health care has improved worldwide during the last 40 years, the incidence of CP is stable and the number of children with more severe forms of CP has even increased. The latter results from the improvement in neonatal-care, which increased survival of the number of extremely premature infants (Baxter et al. 2008).
- The strongest risk factors for developing CP are:
 - prematurity: born before week 28 of gestation,
 - low birth weight (e.g. small for gestational age infants),
 - twins or higher multiple births,
 - perinatal infection,
 - infants that are in poor condition at birth,
 - advanced maternal age and low maternal education attainment,
 - genetic predisposition: autosomal recessive syndromes, usually in association with microcephaly and learning difficulties (e.g. mutation in a Prothrombin gene which confers a mildly increased risk of hemiplegic cerebral palsy).

2.2.2 Aetiology of cerebral palsy

The aetiology of CP should be regarded as a sequence of causal factors occurring in series or in parallel, which ultimately lead to a damaging event(s) to the developing brain. The above-mentioned risk factors are not the causing factors for the brain lesion but may play a variable role in the causal pathway. For instance, prematurity is not the cause of CP, but a preterm infant has a higher risk to experience hypoxia-ischemia, which is causative for the brain damage. The following (series of) pathways are known to be causative for development of CP (Dan et al. 2014; Badawi and Keogh 2013):

- Cerebral malformations (1:100 birth; e.g. microcephaly; pachygyria, lissencephaly). Any malformation of its cortical sensorimotor system will lead to problems in volitional motor function.
- Focal cerebrovascular disease (e.g. perinatal stroke), which can be caused by a blockage (ischemia) or a rupture (haemorrhage) and occur within a timeframe from perinatal until the neonatal period (up to 28 days after delivery). This is the most common cause for unilateral CP in term-born children (Hadzagic-Catibusic et al. 2017).
- Hypoxia-ischemia and cerebral perfusion failure combined or followed by a neonatal encephalopathy, which lead to an increased vulnerability of the grey matter and acute hypoxia-ischemia. This type of injury is responsible for most cases with dyskinetic CP (not for spastic CP).
- Endogenous (e.g. bilirubin) or exogenous neurotoxins, which lead to perinatal brain injury (e.g. kernicterus), often affect the extrapyramidal function as well and lead to dystonia or choreoathetosis type of CP.
- Any prenatal, perinatal/intrapartum or postnatal maternal-foetal infections (e.g. placenta infection), different viruses (e.g. cytomegalovirus which causes multiple intracranial pathologies), bacterial or parasitic infections, which can directly or indirectly contribute to the pathogenesis of CP and are, moreover, common lesions to the white matter.
- Endocrine causes, like the maternal thyroid hormone insufficiency (endemic cretinism), which is known to result in perinatal brain damage in the foetus and results in syndromes that resemble bilateral CP. Postnatal glucocorticoids treatment in preterm infants, which used to prevent chronic lung diseases, have shown to disturb normal brain growth and development as well.

2.2.3 Classification of cerebral palsy

One or a combination of factors listed above can lead to the development of CP. The severity of impairment of an individual child depends on the timing of the injury, the size and location of the brain damage (Baxter 2007), explaining the great heterogeneous clinical experiences within this patient group. To diagnose and help with the classification of these patients, the “Surveillance of Cerebral Palsy in Europe”

(SCPE 2013) offers a decision tree. According to this guideline, a child with CP is characterized by his/her topography of the involved body parts (uni- or bilateral involvement) as well as by his/her predominant tone abnormality (spastic, dyskinetic (with dystonic or chorea athetoid) and ataxic). The use of these classifications systems (Figure 2) allows to furthermore (theoretically) distinguish between the underlying pathology of these tone abnormalities and their motor impairments.

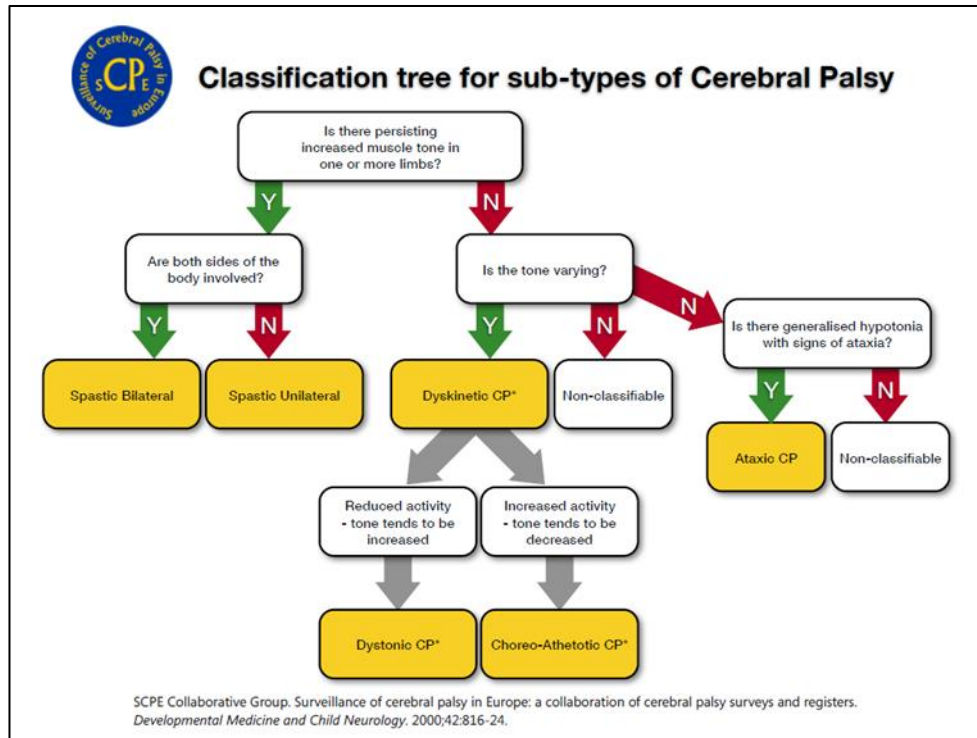


Figure 2: Surveillance of Cerebral Palsy in Europe (2013): Classification tree for sub-types of CP

The spastic group, which is the largest group (70-80%) within CP, is characterized by abnormal movement and/or posture, due to increased tone and pathological reflexes. These symptoms are caused by damage to the white matter (periventricular leukomalacia (PVL)) and are referred to as an UMN lesion. This cortical damage hampers the physiological neural-inhibition via the corticospinal tract (CST), causing an increased neuronal excitability throughout the nervous system (Ludeman et al. 2008; Hoon et al. 2010; Lee et al. 2011; Glenn et al. 2007). Due to this impaired inhibitory function of the pyramidal tracts, the ability of some nerve receptors in the spine to properly receive gamma amino butyric acid is impaired. This leads to an increase of tone (hypertonia) in the muscles innervated by the affected pathways (Fowler 2010). Often, the muscles of the arms and legs are affected. The tongue, mouth and pharynx can be affected as well, which will affect speech, eating,

swallowing and breathing (Kesar et al. 2012; Wittenberg 2009). Spastic CP is described in combination with the topography that describes which limbs are affected, such as spastic unilateral or bilateral CP (Baxter 2007).

The other two forms of CP dyskinetic and ataxic have an extrapyramidal origin, meaning that CNS structures outside the cortex are affected (Bax et al. 2006). This type of CP is present in only in 20-30% of the children with CP. Dyskinetic CP is characterized by abnormal movement and/or posture, which are of involuntary character (uncontrolled, recurring, occasionally stereotyped movements). Pathophysiologically, the nuclei in the basal ganglia are known to be vulnerable to a bilirubin encephalopathy as well as to a hypoxic-ischemic brain injury, mainly causing this type of CP (Rosenbaum 2014). There are two sub-types of dyskinetic CP: dystonia and choreoathetosis. Dystonia is characterised by involuntary muscle activation patterns that result in slow twisting or repetitive movements, or abnormal sustained postures, that are triggered by attempts to move. The term athetosis describes involuntary, slow, continuous movements, which are present at rest and made worse by attempts to move. People with athetosis experience fluctuations in muscle tone – with muscle tone alternating between being floppy (hypotonia) and extremely variable movements (hyperkinesia) making it difficult to maintain a certain posture.

Damage to the cerebellum is the main underlying factor for ataxic CP (Bax et al. 2006). This subgroup is characterized by abnormal movement patterns and/or postures, where movements are carried out with abnormal force, rhythm and accuracy. People with ataxia often experience problems with the balance during walking and with their control of eye movements, eye-hand coordination and speaking.

2.2.4 Underlying pathophysiology of impaired SVMC in children with CP

As explained above, control of movement and/or posture is impaired in all sub-types of CP. Nevertheless, the clinical appearance of impaired motor control, as well as their underlying pathologies differ between sub-types. While children with a lesion to extrapyramidal structures have more problems with the overall coordination and smoothness of movements during rest or while moving (i.e. dyskinesia subtype), children with a primarily pyramidal lesion (i.e. spastic subtype) experience more problems with the selective activation of a specific muscle or muscle group (Rosenbaum 2014; Baxter et al. 2008; Bax et al. 2006). An expert consensus group has defined SVMC as the "ability to isolate the activation of muscles in a selected

pattern in response to demands of a voluntary movement or posture" (Sanger et al. 2006). The term 'voluntary' within SVMC emphasizes the deliberate performance of selected muscle activation during functional tasks (Fowler 2010). As timing, force and speed of voluntary muscle activation are controlled through the CST and its structural and functional connections to other brain areas, damage of the CST is known to causes a loss of selective motor control (Rothwell 1987; Porter and Lemon 1993). How its altered function may negatively influence selective motor control in children with CP will be described and discussed in detail below. In the next section, the following four main points will be discussed with regard to the pathophysiology of SVMC: i) development and maturation of the CST; ii) motor cortex mapping; iii) CST networking and iv) sensory input and feedback.

2.2.4.1 Maturation, functioning and impairment of the corticospinal tract

The CST is a white matter motor pathway, which originates in the cerebral cortex and descends to alpha motor neurons and interneurons in the spinal cord. Neurons of the CST are UMNs that directly control muscle activation and thereby movement control. Structurally, the CTS consists of different parts: the crossed lateral CST, the uncrossed lateral CST, and the uncrossed anterior (ventral) CST. Some authors even report a fourth, the crossed anterior CST (Nathan, Smith and Deacon 1990). Although the functions of these different parts of the CST are known to differ, the exact functional role of each CST is still discussed and investigated. Overall, the CST is associated with the acquisition of (dexterous) motor skills. As the crossed lateral CST occupies 75–90% of the CST fibres, extending caudally to the dorsolateral fasciculus to the last sacral segment after crossing the medulla, it is the largest tract that iscontrolling mainly distal muscles (i.e. finger, ankle) involved in fine motor movements. Proximal muscles of the upper extremity as well as trunk and neck muscle are thought to be controlled by the uncrossed anterior CST (5-15% fibres of the CST) and other neural tracts, such as the cortico-reticulospinal tract (Jang 2014). Additionally, there is some evidence that together with the cortico-reticular tract, the anterior CST is involved in walking and postural control (Jang et al. 2013, Barthélemy et al. 2011). The function of the uncrossed anterior CST, which descends within the lateral funiculus and is located ventrally to the crossed lateral CST, has been suggested as a motor recovery pathway (Jang 2014). Following an injury of the lateral CST, it functions as an ipsilateral motor pathway from the unaffected motor cortex to the affected extremities together with the cortico-reticulospinal tract (Jang et al. 2013;

Staudt 2007). Growth and development of the CST tract within the spinal cord, as well as maturation of connectivity of these neurological structures, are explained in the next paragraph.

a) Development and maturation of the CST

Before the first CST axons invade the spinal cord, proprioceptive axons reach the intermediate grey matter of the spinal cord. This occurs at 7.5 weeks post-conception (Clowry 2007). Next, the motor neurons are innervated around 8.5-9.5-week post-conception. This afferent ingrowth and increased density of axodendritic synapses can be observed intrauterine in the foetus, as movement shifts from spontaneous holokinetic movements into stimulus-evoked ideokinetic movements around the 8.5 weeks post-conception (Welniarz, Dusart and Roze 2017). With full innervation of the ventral horn, the monosynaptic reflex arc is established around 14 weeks post-conception. After muscle afferent input is completed, CST axon projection from the cortex to the spinal cords begins around the 17th-week post-conception. First CST fibres invade the grey matter around weeks 20 to 27, and fully innervate the ventral horn until the 35th-week post-conception (Welniarz, Dusart and Roze 2017). At the age of 2 years and under normal development, the synaptic network of the CST in the spinal cord with alpha motor neuron and the interneurons, as well as myelination of the tract is completed (Welniarz, Dusart and Roze 2017; Jang 2014; Clowry 2007). During this time, the CST and spinal cord undergo refinement of connectivity, eliminating polyneuronal innervations of muscles, into more selective ones. This refinement involves subsequent synaptic withdrawal as well as strengthening of repetitively activated connections (Welniarz, Dusart and Roze 2017; Clowry 2007). Selectivity of this neural refinement is an activity-dependent neuroplastic process, which involves structural as well as neurotransmitter modulation (Clowry 2007). During this critical developmental period of CST maturation, a high activation of N-methyl-D-aspartate receptors has been observed, which regulates dendrites growth as well as it eliminates gap junctions between motor neurons. CST ingrowth further triggers upregulation of parvalbumin expression, which promotes inhibitory synapse formation. Due to these processes' afferent, input into the ventral horn is subsequently reduced and inappropriate heteronymous connections are further reduced (Clowry 2007). This improved neural refinement can clinically be observed, for instance in the occurrence of more mature (slower onset and greater threshold) stretch reflex response, upper-extremity fine-motor movements and independent walking in one-

year-old toddlers (Welniarz, Dusart and Roze 2017; Clowry 2007). If, however, CST damage occurs within in this critical period of developmental, the segmental and circuitry refinement is disturbed, which is known to alter motor development and control permanently (Welniarz, Dusart and Roze 2017; Clowry 2007; Staudt 2007). A developmental lesion of the cortex will alter innervation as well as gene expression of the spinal cord, resulting in a large number of muscle afferents in the ventral horn. In the activity-dependent competition for synaptic spaces within the spinal cord, physiological elimination of these muscle afferents is hindered by altered cortical activation in favour of more appropriate synaptic formation. This pathophysiological dominance of muscle afferents in synaptic space has been shown in several animal studies (Clowry 2007). These studies showed that unilateral inhibition of one cortex during this critical developmental period resulted in less active projection of the injured cortex, as well as it increased inappropriate activity from the ipsilateral spared cortex. When on the other hand muscle afferent activity was reduced, for instance by the means of BTX injection, studies in kittens showed that this resulted in fewer contralateral synaptic connections with CST axons (Martin et al. 2004). This suggests that synaptic refinement within the spinal cord can only happen when both segmental (peripheral) and descending (central) pathways serve to reinforce connectivity. Applying this finding to children with cerebral palsy explains why gaining adequate motor control is more difficult for them as compared to their normally developing peers. Firstly, due to their brain lesion, the output of their descending pathways is reduced, which in turn results in diminished movements (Staudt et al. 2003). Secondly, this diminished CST output results in non-physiological movement as well as in overall reduced motor activity and thereby in decreased segmental input (proprioception) required for refinement of synaptic connections. Furthermore, as the brain damage in children with CP often also affects the developing brain, their movement might be altered and reduced already intra uterine (Campbell et al. 2018). The overall effect is a distributed neuro-motor-control-circuit from cortical output and input. For instance, children with prenatal brain lesions show less frequent and more mass movements intrauterine (Horimoto et al. 1993). Postnatally, within the first month of development, this can further be observed by a paucity of leg movements, prolonged, monotonous kicking and less purposeful leg control (Campbell et al. 2018; Heathcock et al. 2005; Fetters et al. 2004). Another study comparing the acquisition of motor control between infant born preterm and term-born infants showed that preterm born infants require significantly more time to learn that their kicking can affect

movement of a mobile-play (Heathcock et al. 2004). Furthermore, the time window in which the decoupling of the movement of intra-limb joints is learned and which is a prerequisite for more complex purposeful movements, is prolonged in pre-term infants (Heathcock et al. 2004).

b) Motor cortex mapping

Early during brain development, corticospinal neurons can be found throughout the frontal, parietal, occipital and temporal lobes. However, through the development and refinement process of CST and spinal cord described above, the distribution of these neurons becomes restricted to the posterior frontal and anterior parietal lobes (Nathan, Smith and Deacon 1990). These cortical areas are associated with the primary motor cortex (M1), secondary motor area and somatosensory cortex.

The understanding and definition of the different motor areas within the motor cortex started in 1870 by electrical stimulation experiments from Hitzig and Fritsch. They demonstrated that exciting the motor cortex in a dog could evoke muscle twitches. Within the same year, Ferrier did similar experiments with monkeys. In contrast to Hitzig and Fritsch, these experiments resulted not only in muscle twitches but also in more complex body movements. Ferrier concluded that the motor cortex controls complex features of movement rather than muscles per se, as Hitzig and Fritsch were suggesting (Rothwell 1987). Since then, these controversial findings have been discussed by many other researches within the field and are often phrased as the "muscles versus movement" debate (Philips 1975). One explanation for the diversity in findings is related to the difference in the electrical stimulations performed in these experiments: Stimulation of almost any area of the cortex (with sufficient intensity and under the appropriate conditions) can evoke complex movement. Although today, the true mosaic representation of CST, which favours a very selective muscle activation (i.e. comparing the motor cortex with the keys of a piano), is replaced by a notion of overlapping organization of CST cells, consensus with regard to the CST organization is still lacking. For instance, one group (Asanuma, Larsen and Yumiya 1980) suggests that the cortical neurons controlling a single muscle are focally arranged in a so-called interconnected cortical "efferent zone" in M1. In contrast, Phillips and Porter (1977) speak of CST neuron "colonies" for different muscles, which can overlap extensively within this cortex and form a 'discrete' or 'mosaic' fine-grained structure of cortico-motor neuronal output (Omrani et al. 2017; Rathelot and Strick 2006). Others argue

that the primary motor cortex should be regarded as a dynamic machine, which does not directly code for movement parameters, but instead should be understood in terms of the rules that govern its pattern generation (Kaufmann et al. 2016; Graziano et al. 2002).

The primary motor area of this so-called M1 was found to be the largest of those three. Anatomically, it is laying along the precentral gyrus (Rothwell 1987). According to its map, movement control of the lower extremities is activated by the most medial parts, whereas movement of the upper extremity and face are located in the most lateral parts (Fig 3). In addition to this so-called medio-lateral projection, there is also a rostral-caudal gradient, as movements of the distal body parts are nearest to the central sulcus, whereas proximal body parts are represented more rostrally. Beside the different M1 locations, the relative size of areas devoted to movement of particular body parts differs. For instance, cortical areas of distal and fine-tuned movements (i.e. finger, hand, tongue) have a much larger representation than those of body parts of more proximal and gross motor movements. In order to selectively control these distal joint movements more, cortical output cells to the muscles and the local interneurons are needed, thereby engrossing a larger area. Graphically, this phenomenon is visualized by the so-called homunculus (Rothwell 1987).

Next to M1, the most lateral part of the precentral gyrus, the secondary motor area (M2), is extending into the lib of the Sylvian fissure. Body representation of the M2 is the reserved version of M1. Compared to M1, M2 has been investigated less and its exact functioning is still being explored. Today, its role is mainly described in linking sensory cues to motor actions and thereby enabling adaptive choice behaviour (Barthas and Kwan 2017).

The third motor area, which is termed the supplementary area (SMA), is rather than M3 found on the medial surface of the hemispheres. It differs from M1 in four main aspects: i) its threshold for stimulation is higher, ii) its somatotopic representation is poor; iii) its activation results in eliciting complex multi-muscle and –joint movements as well as iv) its activation is mainly bilateral. Functionally, it is therefore associated with movement's maintaining postural control and closely connected to other brain areas (Ruan et al 2018).

In conclusion, due to the functional complexity of even normal motor control, a prediction of the nature of a certain impaired motor outcome by solely identifying the impaired underlying brain structure (somatotopic) is too simple.

Nevertheless, the general structure and function of the motor cortex as described above, might help to explain certain clinical symptoms, for instance the occurrence of a phenomena called “proximal-distal-concordance”. This describes the increased distal impairment of lower extremity motor function in comparison to more proximal joints. As this phenomenon is mainly observed in children with spastic CP, it is thought to be related to impaired CST functioning following a PVL (Baxter 2007; Bax et al. 2006). The PVL is characterized by multifocal cortical white matter necrosis and often occurs during the period with the greatest vulnerability of the brain somewhere between 26 and 34 weeks of gestation (Beaulieu and Schneider 2013; Rha et al. 2012; Lee et al. 2011; Thomas et al. 2005; Staudt et al. 2003). As the CST fibres supplying the lower extremities and distal joints lie closest to the periventricular area (Fig 3), these fibres are more likely to be damaged than those providing more proximal joints (Staudt et al. 2003; Glenn et al. 2007).

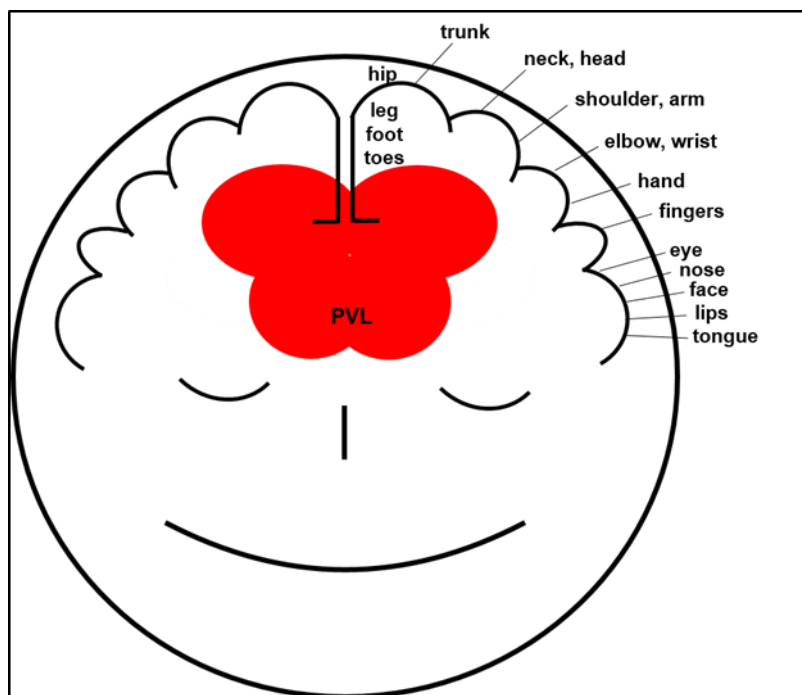


Figure 3: Periventricular Leukomalacia and somatotopic organisation of the motor cortex, showing the close proximity of the lesion, CST and motor cortex

c) CST Networking

As discussed above, movement control cannot sufficiently be predicted by motor cortex representation (e.g., homunculus) alone. Although the CST itself might be considered a critical structure for motor control, research has shown that impaired motor control does not only occur due to a pure lesion of this structure alone (which is quite rare), but rather due to impaired CNS networking (Rothwell 1987). For instance, shortly after the experimental bilateral pyramidotomy in monkeys, these animals show normal gross motor movements (sitting, running, climbing, head control) (Lawrence and Kuypers 1986). They only experience problems in fine motor control of the fingers and hand (i.e. pinch grip for picking up food). Structurally, this explanation is reflected in the presence of direct monosynaptic (cortico motor neuron) projections from the CST, as well as polysynaptic ones and its collaterals to other structures of the central nervous system (i.e. brain stem, cerebellum, spinal cord). This network action or in-concert action of the CST with other brain structures is also evident in humans and will be described in the next two sections in relation to studies on gait control and the disappearance of mirror movements.

Studies investigating motor control of walking in acephalic children have shown that the CST is not responsible for initiating and activating synergistic stepping patterns but might be more accountable for fine-controlling joint movements, especially around distal joints. This observation shows that intact midbrain structures, which are the origin of the rubrospinal tract (RST), seem to be responsible for the activation of a more stereotyped stepping movement (Hicks and Onodera 2012; Yeo et al. 2012; Yang et al. 2011). If both tracts are working functionally together under normal physiological conditions, the fine-tuning of the walking pattern is possible.

Regarding the natural maturation process of learning to suppress mirror movements (i.e. simultaneous movement of the contralateral joint), studies have shown that adequate cortico-cortical connections (networking) are required. Mirror movements can be observed in young children (under the age of 10 years), when asked to perform difficult movements. Neurophysiologically, the interhemispheric inhibition between the two M1s via the corpus callosum has not yet been sufficiently established in young children and thus allows bilateral cortical activation. As this refinement process is naturally reinforced by development in combination with repetitive activation (activity-dependency) (Koerte et al. 2010), the degree to which ipsilateral projections disappear in children with CP depends on the time point the damage occurs and the

severity of the damage. In general, early damage and largely disturbed motor output is associated with continuing ipsilateral projection. If the damage occurs later during development, some ipsilateral projection might be eliminated but other act together with contralateral ones, thus resulting in a mixed picture of ipsi- and contralateral organization (Kuo, Friel and Gordon 2018).

In summary, motor control dysfunction in children with CP occurs not solely due to the white matter lesion of the CST itself, but even more profoundly due to its descending and ascending connections to several neural areas. Such as i) connections between different cortical motor and sensory areas, ii) the main motor output pathway to the spinal cord, iii) the loops connecting the cortex with the basal ganglia and cerebellum; and iv) ascending sensory information from the thalamus to the sensorimotor cortex (Forssberg 2014; Glenn et al. 2007; Lauer et al. 2005; Staudt et al. 2003). In relation to the disturbed descending pathways, cortical output cannot reach the motor neurons in the spinal cord and further disturb the typical development of all neural loops, resulting in secondary alternations of muscle activation, growth, and development (Cahill-Rowley and Rose 2014; Rose and McGill 2005). As the functional connection of cortical and subcortical pathways are working together, motor dysfunction in children with CP can affect many different systems (Staudt et al. 2003), which will be described in more detail in chapter 2.2.7 on impaired SVMC in children with CP along the ICF continuum.

d) CST sensory input and feedback systems

Another fundamental prerequisite of motor control is the integrity of the sensory feedback pathways. Any damage to the integrity of the sensory feedback system inevitably causes motor output as well as motor learning deficits (Marsden 1998; Sarlegna et al. 2006). There is an extensive network of direct (i.e. primary somatosensory cortex) and indirect pathways (i.e. ventral posterior nucleus of the thalamus) that deliver peripheral sensory information to the primary motor cortex. Integrating the sensory feedback system in our understanding of motor control increases its complexity exponentially. In the context of this thesis, sensory input in children with CP is often diminished. Wingert and colleagues have shown that diplegic or hemiplegic CP can have proprioception deficits in all limbs (Wingert et al. 2009). This in turn alters their cortical input and thereby the output of the circular system. As the aim of this PhD lies in determining motor control output, there will be no further

detailed explanation of the structural and functional networks of the sensory system with the CST.

In summary, this section provided an overview of the role of the CST in initiating and controlling SVMC. By describing its activity-dependent development and synaptic ingrowth into adjacent central new structures, forming functional networks, it was demonstrated that impaired CST functioning results in decreased descending motor output and in the disturbed functioning of related structures. As the organization of the CST networks is highly complex, clinical symptoms of CST lesion are known to vary in appearance and intensity. The clinical symptoms occurring due to an injury to the CST and resulting impact on the connection to other structures of the central nervous system (as described in section 2.2.4.1) are called UMN signs (Sanger et al. 2006).

2.2.5 Clinical symptoms of corticospinal tract lesions

The clinical symptoms occurring due to an injury to the CST and resulting impact on the connection to other structures of the central nervous system (as described in section 2.2.4.1) are called UMN signs (Sanger et al. 2006). They can be either so-called “positive UMN signs”, such as spasticity, hyperactive reflexes, abnormal posture, dyskinesia, persisting developmental reactions and typically develop due to the missing inhibitory control of the CST (Sanger et al. 2003). Impaired SVMC, muscle weakness, as well as impaired coordination, belong to the negative UMN signs (Sanger et al. 2006). These negative signs are more difficult to observe and to measure clinically. Consequently, the focus of research and medical treatment rested on the assessment and treatment of the positive UMN signs, for instance spasticity for a long time (Cahill-Rowley and Rose 2014; Dobson 2010). However, the influence of negative UMN signs has revealed to be as limiting, or even more limiting for motor performance in children with CP. An increased scientific interest for these impairments has arisen within the last 15 years (Vos et al. 2016; Park and Kim 2013; Voorman et al. 2007; Desloovere et al. 2006; Østensjø et al. 2004). The difficulty of investigating positive and negative UMN signs separately lies in their common pathogenesis and their interconnected hidden neural pathways. Therefore, clinically, children with CP often present both positive and negative UMN signs (Figure 4) (Carr 2014). For example, a child with spastic (positive UMN sign) bilateral CP will also suffer from impaired SVMC and muscle weakness (negative UMN signs).

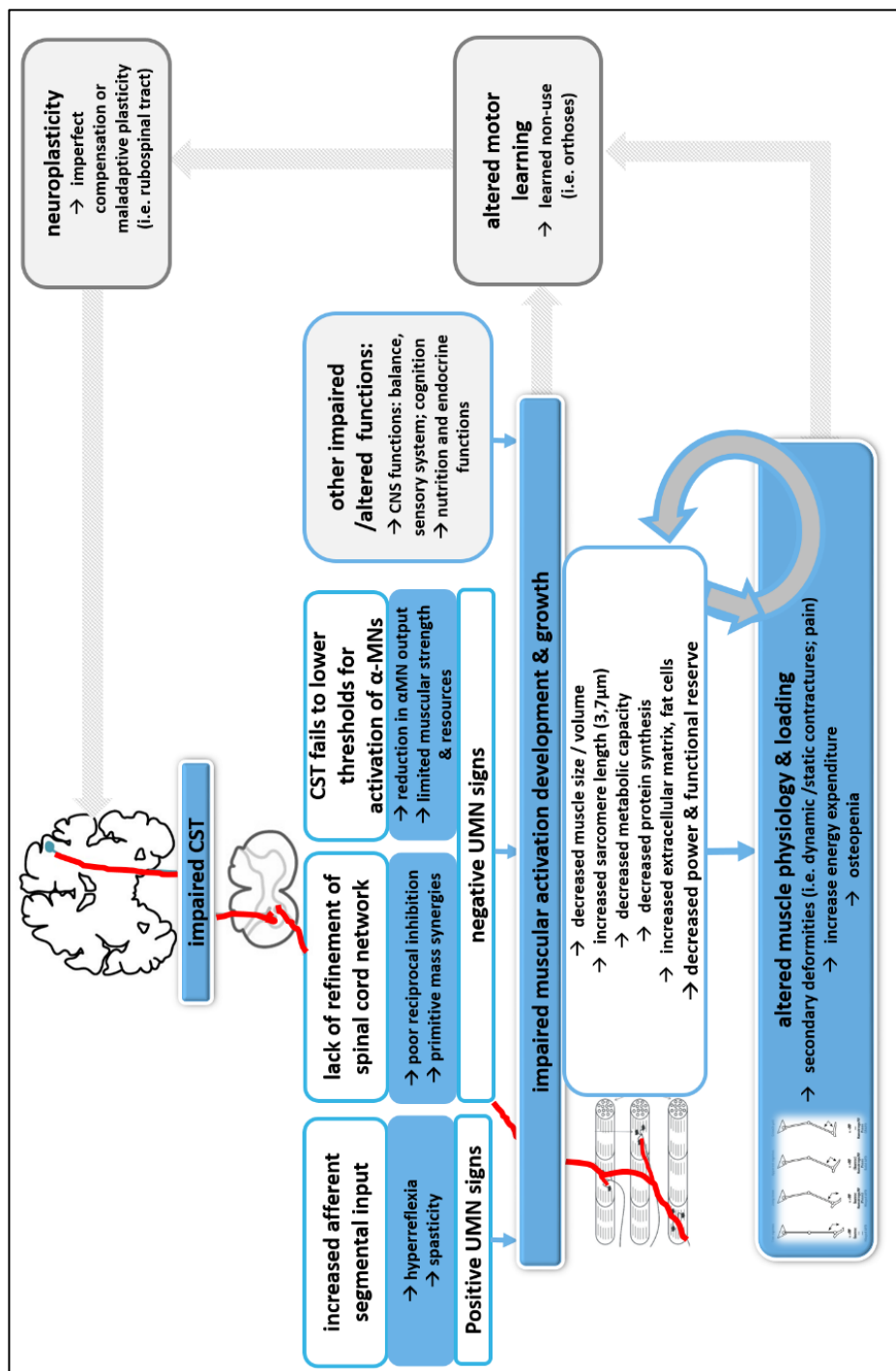


Figure 4: Possible symptoms due to an UMN lesion: showing its complexity and long-term vicious circle effect in children with CP.

Starting with a lesion at the CST, which disturbs efferent output at the level of the spinal cord and brain centres, it causes positive and negative UMN signs to appear. These positive and negative UMN signs themselves and in combination with altered endocrine and nutrition functions are then causative for impaired muscle activation and growth, which in the long term leads to altered muscle physiology and loading. In relation to these impaired control and muscle prerequisites, a child with a UMN lesion learns to move in a different way and possibly also with the help of other brain structures and tracts (e.g. rubrospinal tract) within the context of neuroplasticity.

The following three neurophysiological processes which are caused by an injury to the CST are responsible for this mixed presentation of positive and negative UMN signs. First, due to impairment of the CST, the afferent segmental input is increased, which causes symptoms like hyperreflexia and spasticity (Hadzagic-Catibusic et al. 2017; Hurvitz et al. 2014; Fowler 2010). As spasticity mainly involves antigravity muscles (e.g. extensor and adductor muscles in the lower extremity), an imbalance of muscle tone between the agonist and antagonist of the joint develops (Fowler 2010). This can further lead to shortening of the hypertonic muscle around the joint and limits its range of motion or makes moving the joint painful.

Second, a lack of refinement of the spinal cord networks disturbs reciprocal inhibition and causes primitive movement patterns to maintain (Hurvitz et al. 2014; Fowler 2010). As the process of reciprocal inhibition is mainly associated with the appearance of impaired SVMC, this will be discussed in more detail in section 2.2.6 ("Altered processes, structures and networks in relation to impaired SVMC"), together with other involved neuroanatomical structures and networks.

Third, the impaired cortical spinal control fails to lower thresholds for activation of alpha motor neurons, which results in a reduction in alpha motor neuron output and is thereby limiting the ability to activate the muscular resources fully. In the long term and especially during development, these three commonly altered neurophysiological processes due to UMN lesion alter normal muscle development and growth in children with cerebral palsy. Research on the muscle-morphology, -physiology and -metabolism of children with spastic CP has shown the following changes due to the diminished central control: i) increased resting sarcomere length (3,7 μm), which alters the muscles position of optimal force production, and, therefore, might be one mechanical cause for the appearance of (dynamic) contracture (Smith et al. 2011; Shortland 2017); ii) decrease in muscle volume about 1/3 in comparison to neurological intact children (Vanmechelen et al. 2018; Noble et al. 2017; Noble et al. 2014; Shortland 2011); iii) reduced elasticity (active component) and muscle-belly length and size (Barber et al. 2011); iv) atrophic muscle cells; v) increased amount of extracellular matrix (Noble et al. 2014); vi) reduced amount of satellite cells, which results in a limited ability to repair muscle (Smith et al. 2013); vii) reduced amount of type II muscle fibres (Hurvitz et al. 2014; Petersen et al. 2013; Downing et al. 2009; Poon and Hui-Chan 2009). Consequently, the muscle strength in children with CP is reduced compared to neurologically intact children (Shortland 2011). Altered

nutritional and endocrine factors, which are commonly found in these children as well, additionally worsen this muscular situation, especially regarding muscular synthesis and repair. All these altered muscular requisites lead to an altered muscle loading during function and activity and can furthermore lead to the development of musculoskeletal malformations, osteopenia, injury and chronic pain (Shortland 2011; Peacock 2009).

2.2.6 Altered processes, structures and networks in relation to impaired SVMC

Normally, when a muscle spindle is stretched, its stretch reflex is activated, and the opposing muscle group will be inhibited to prevent it from working against the resulting contraction of the homonymous muscle. This inhibition is controlled by the cortical motor output to inhibitory interneurons in the spinal cord. This inhibition prevents the opposing alpha motor neuron from firing, thereby reducing the contraction of the opposing muscle. Without this reciprocal inhibition, agonistic and antagonistic muscle groups might contract simultaneously (e.g. an undesired co-contraction), inducing a co-contraction and thereby increasing the net force on the joint. This finally results in higher energy expenditure. Mostly synergistic movement flexor/extensor patterns occur, which are normally suppressed under cortical control. The appearance of these primitive movement patterns associated with impaired SVMC is considered to result mainly due to the altered control of extrapyramidal motor tracts that originate in the brainstem (rubospinal, reticulospinal and vestibulospinal tracts) and central pattern generators (CPGs), which are located in the spinal cord. Especially the rubospinal tract (RST) seems to mediate gross movement patterns, those promoting extension and inhibiting flexion (Cahill-Rowley and Rose 2014; Fowler 2010). This function was shown when observing movement in anencephalic children. These children can perform synergetic stepping patterns, although they are born with no cortices (Peiper 1963). Therefore, their intact midbrain structures, origin of the RST, seem to be responsible for the activation of these stereotyped movements (Hicks and Onodera 2012; Yeo et al. 2012; Yang et al. 2011). Furthermore, the RST has found to be more anatomically and physiological prominent during early human development, which might explain why infant movements rely more on co-activation of synergists and antagonists than adult movements (Forssberg 1985). It is suggested that human movement becomes more and more selective by the developmental transition away

from RST reliance to primarily CST reliance (Cahill-Rowley and Rose 2014; Porter and Lemon 1993; Forssberg 1985).

Other white matter structures which have also shown to play a role within diminished SVMC are the transcallosal motor fibres. As these fibres connect the corpus callosum and the primary motor cortices of the two hemispheres, they are important for the process of interhemispheric inhibition (Forssberg 2014; Kwon et al. 2014; Lee et al. 2011; Lebel et al. 2012). Interhemispheric inhibition increases during normal development (Lebel et al. 2012). It enables a person to perform reciprocal alternating as well as selective movements, and it decreases the occurrence of involuntarily associated movements (Forssberg 2014; Kwon et al. 2014; Lee et al. 2011).

Furthermore, the functioning of the sensory system is essential for SVMC. Besides the recruitment of the correct number of muscle fibres per motor unit, adequate proprioception is essential for SVMC (Rothwell 1987; Miller 2006). Especially sensory feedback is important when learning new voluntary movement strategies. Motor learning research has shown that general motor programs for basic/innate motor behaviours (e.g. eating, walking) are located within the spinal cord (Forssberg 2014). The neural networks, which are responsible for activating the “stored” movement programs involving multiple joints, are called CPGs (Barthélemy et al. 2011). The CPGs are controlled by descending systems and form the cortical and subcortical centres. As this cortical control is diminished in children with CP, their ability to initiate, adapt and stop their stored motor programs, for instance for learning to walk, is diminished (Begnoche et al. 2016). Recent studies on human walking have shown that especially the control via the CST plays an essential role in adapting the basic locomotor rhythmical movements such as during gait. For instance, a study by Barthélemy and colleagues (2011) reported correlations between a reduced corticospinal excitability of the tibialis anterior and the degree of foot drop during gait in patients with spinal cord injuries. Therefore, although basic motor programs are stored within the CPGs, their performance and their performance quality depend on cortical structures (Barthélemy et al. 2011).

In summary, the aetiology of impaired SVMC is still in its infancy. Evidence suggests that the CST plays an essential role due to its mono- and polysynaptic connection with other areas central nervous structures (e.g. basal ganglia, spinal cord (CPGs)). Impaired SVMC could be caused by a loss of connections to the CST and thus would be classified as a negative UMN sign, a loss of inhibition of primitive flexor/extensor

patterns, which would be a positive UMN feature, or by a combination of both (Dobson 2010). In relation to the complex pathophysiology of UMN signs, evaluating whether a lack of muscle force is primarily caused by a lack of strength, or if this weakness is more related to the inability of selective muscle activation, is not always clear. Similarly, establishing whether hypertonia is hampering movement, SVMC, or both, is challenging but of importance for the identification of an appropriate interventional approach.

2.2.7 Impaired SVMC in children with CP along the ICF continuum

The ICF continuum will be used as a common framework (Rosenbaum and Stewart 2004) to describe the clinical presentation and functional consequences of impaired SVMC of the lower extremity in children with CP. The ICF was developed by the World Health Organization to describe the health status and the consequences of any illness for each individual person in a holistic manner. It presents the person's health status on different levels: body structure and function level (health condition: e.g. PVL), activity level (functional consequences: e.g. gait problems), participation level (social consequences: e.g. attending a special school) and furthermore important personal and environmental factors (Figure 5). Impaired SVMC is listed in the ICF- Children and Youth (ICF-CY) under body functions "b7600: Control of voluntary movement functions - Functions associated with control over and coordination of voluntary movements" (Schiariti and Masse 2014). Although one could say that SVMC might not be of great importance for the child's quality of life, as long as he/she can carry out all desired activities (e.g. walking), this statement might be only partly true. Impaired selective activation can initiate and worsen a vicious cycle of limited active movement, joint contractures, hampered motor function and diminished activity and participation (Levin et al. 2009; Wren et al. 2005; Bell et al. 2002; Johnson et al. 1997) (Figure 6). How this vicious cycle can be initiated by impaired SVMC and how impaired SVMC might thereby have consequences/influence on all ICF levels will be described in the next sections 3.2.7.1 – 3.2.7.3.

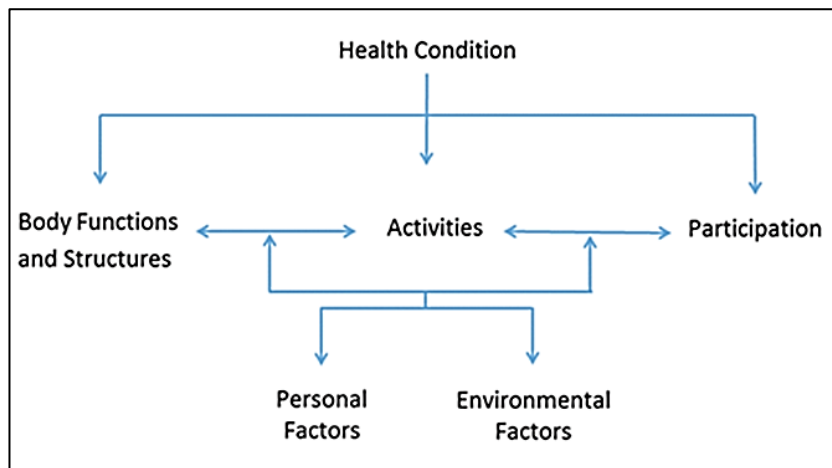


Figure 5: International Classification of Functioning, Disability and Health (ICF) continuum:

Showing all levels which influence health, body function and structure, activity and participation level as well as personal and environmental factors and their interdependency.

2.2.7.1 SVMC impairments on body functions and structures level

As explained in the section above, the underlying injured body structure in children with impaired SVMC is the motor cortex and the CST, often due to a PVL (health condition). Impaired SVMC is known to result in the following two impairments on the body function level: the appearance of primitive mass synergy patterns and/or mirror movements (Cahill-Rowley and Rose 2014). Historically, Brunnstrom was the first who described the emergence of mass flexor and extensor synergies during motor recovery (stage 2) from hemiplegic stroke (Huang et al. 2016). This observation in stroke patients was supported by Perry (Perry 1993) who also observed these primitive patterns of mass extension and flexion in children with hemiplegic CP. In contrast to the adult stroke population, the appearance of mass synergy patterns in children with CP is also caused by the impaired maturation of the CNS's structures and functions (Levin and Panturin 2011; Levin et al. 2009). For instance, during normal human motor development, tightly coupled stereotyped kicking movements of the lower extremity can be observed during the first month of living (Fetters et al. 2004). With increased maturation, involving myelination of the CST, activation of the hip, knee and ankle joint becomes more and more uncoupled. This maturation process is impaired in children with CP. Therefore, an abnormal coupled kicking pattern persists in these children and will furthermore hamper and delay the child's ability to reach other motor development milestones (e.g. crawling, walking) (Mayston 2001). Another possible explanation for the appearance of mass synergies is an

increased cortical overlap of joint representation, which has been found for the upper extremity presentation in hemiparetic stroke patients.

Maturation of the CST includes increased pruning of the axons. Its initial bilateral projection changes to contralateral within the first year of living (Van de Winckel et al. 2013; Lebel et al. 2012; Eyre et al. 2001). Consequently, SVMC of the right lower extremity is controlled by the left motor cortex. Observations of children with unilateral spastic CP have shown that this change in cortical projection is use-dependent. Other investigations (Carr et al. 1993; Staudt et al. 2003) have shown that, depending on timing (prenatal, around birth and postnatal lesion) and extent of the lesion, children will have either contralateral (normal) projections (small lesions), both contralateral and ipsilateral projections (medium lesions) or ipsilateral projections (large lesions). Clinically, mirror movements often occur from the persistence of ipsilateral corticospinal projections between the non-lesion motor cortex and the paretic limb. Mirroring is defined as simultaneous, obligatory movement of the contralateral limb during ipsilateral movement of the same joint (Fowler 2010).

2.2.7.2 Possible SVMC caused limitations on the activity level

To show the connection between SVMC and possible limitations on the activity and participation level (section 3.3.7.3), back referencing to the body functions and structures level was necessary.

At the ICF activity level, impaired SVMC interferes with motor performance, as it was recently shown by different studies. Some of these studies even suggest the superior importance of impaired SVMC in comparison to other UMN impairments.

One possible consequence of impaired SVMC is that children with CP have difficulties in learning new motor skills and thereby reaching motoric milestones such as walking much later than their peers. The majority of children with milder forms of CP (e.g. Gross Motor Function Classification System (GMFCS) level I and II), who learn basic gross motor skills within their first 2 years (e.g. to roll from a supine to a prone position and sit without support), walk independently between 3 and 5 years of age. Children with more severe forms of CP are functionally classified as GMFCS level III to V and do not learn to walk independently at all. These children either need a hand-held mobility device for walking (GMFCS level III), powered mobility devices (GMFCS level IV) or the help of another person for transport (GMFCS level V). The ability for reciprocal lower limb movements (Fedrizzi et al. 2000) is one of the most important

predictors of independent walking in children with CP. Reciprocal lower limb movements describe alternating limb movements to advance the body forward through crawling, walking and running. This coordinative process is neurologically interlinked with the maturation of white matter structures like the CST and the corpus callosum (transcallosal motor fibres), which are also involved in the process of interhemispheric inhibition (Begnoche et al. 2016; Peacock 2009; Gage et al. 2009).

One of the most apparent characteristics of many ambulatory children with CP is that they do not walk 'normally', e.g. with different gait patterns in comparison with typically developing children (Gage et al. 2009; Miller 2007). This often means that their walking endurance and/or capacity to perform other tasks while walking (dual-tasking) is often reduced compared to their healthy peers (Postans and Granat 1999; Begnoche et al. 2016; Meyns et al. 2012). One of the reasons for the altered gait pattern is the decreased ability to achieve a physiological limb position in several phases of the gait cycle caused by extensor synergy patterns in the lower extremity. For instance, during the late swing phase, the physiological limb position is characterised by selective knee extension while hip and ankle remain flexed (Gage et al. 2009). In children with impaired SVMC, this uncoupled activation is replaced by an extensor-synergy pattern (obligatory co-activation of the quadriceps and gastrocnemius) (Zwaan et al. 2012; Damiano et al. 2000). As such overlap in muscle activation does neither occur in the neurologically intact children nor in children with idiopathic toe walking, this co-activation has been attributed to the neuropathological nature of impaired SVMC (Cahill-Rowley and Rose 2014). This synergistic extension pattern is one possible cause for forefoot contact during initial contact, which further hampers physiological weight bearing during stance and increases the risk of trips and falls (Rha et al. 2016).

2.2.7.3 Possible SVMC caused disabilities on participation level

The walking limitations described above, which are caused by impaired SVMC as well as by other common impairments (e.g. lack of muscle strength, balance, endurance), have shown to be potential barriers to participate in physical, recreational and social activities in the life of children with CP. A study by Palisano and colleagues (Palisano et al. 2004) found that children classified at GMFCS levels II and III were 4.6 times more likely to not participate in any activities with friends or others compared with those who walked without assistance (e.g. GMFCS level I). Therefore, the ability to walk without support, even for short distances, may affect independence in mobility and

participation in recreational, leisure and learning activities of the child and its family. Although the GMFCS level does not describe the quality of the gait pattern, lower GMFCS levels (I and II) are associated with more efficient walking patterns, which are known to be associated with more normal activation patterns. A study from Riad and colleges (Riad et al. 2013) showed that decreased selective motor control during walking has an influence on self-esteem (ICF personal factors) in children with mild unilateral spastic CP. Although these children are physically high functioning (GMFCS level I), their gait pattern deviations (Gait Profile Score) correlated with lower self-esteem. This was even more pronounced when non-physiological arm movement was involved.

To conclude, although impaired SVMC is an impairment on the body functions and structure level, it might cause, together with other common UMN signs, limitations and disabilities for children with CP and thereby reducing the child's and family's overall quality of life (Figure 6).

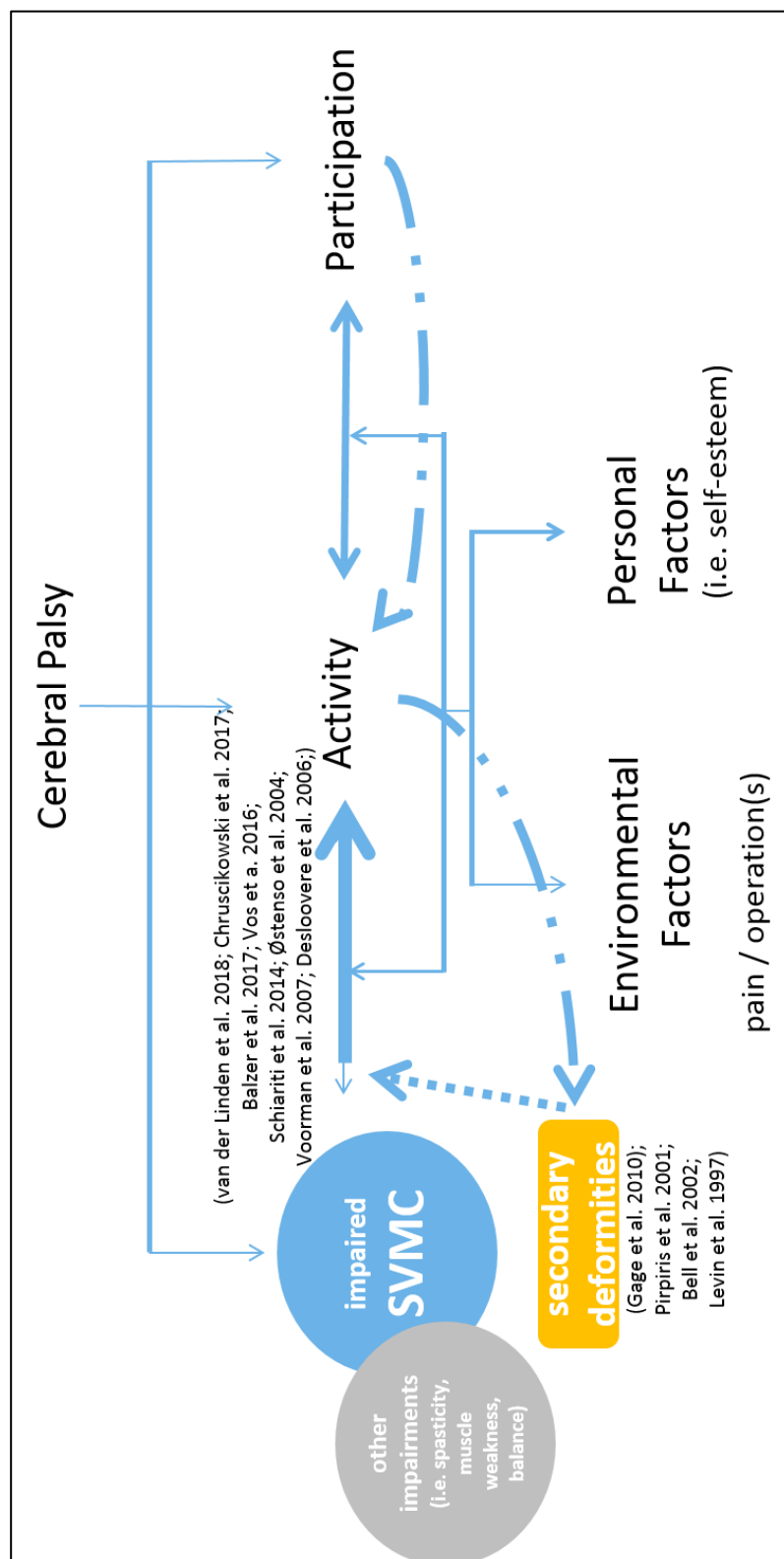


Figure 6: Clinical consequences of impaired SVMC presented at the ICF levels: Impaired SVMC has shown to negatively influence the child's activity as well as its participation level. Impaired SVMC may also cause secondary deformities in the long term, which might further lead to pain and/or surgical correction.

2.2.8 Trainability of SVMC

The training to move as physiologically as possible is one of the central physiotherapeutic goals within neurological rehabilitation (Levin and Panturin 2011; Miller 2007; Mayston 2001). The rationale behind this therapy focus is to improve the patient's movement efficiency as well as to reduce the development of secondary impairments like contractures and muscle insufficiency (Mayston 2014; Gage et al. 2009). As most medical treatments and surgeries are focusing on reducing positive UMN signs like spasticity, physiotherapy and occupational therapy, they are the only disciplines that are trying to influence negative UMN signs like muscle weakness and impaired SVMC (Dobson 2010; Fowler 2010). A variety of therapy concepts exist (e.g. Neurodevelopmental treatment approach (NDT) / Bobath concept; Brunnstrom) aiming to improve voluntary, physiological activation of muscles in neurological patients (Levin and Panturin 2011; Pandian et al. 2012; Barber 2008; Mayston 2001). These approaches are based on the theories of motor learning and neural plasticity. These concepts suggest that the (re-)training of physiological movements is possible via the adaptive capacity/reorganization of the CNS (Shishov et al. 2017; Forssberg 2014; Mayston 2014). Fundamental for this neural-plasticity is that sensorimotor learning takes place in a functional, motivating, repetitive and possibly adaptable environment (Johnston 2009; Trojan and Pokorný 1999). Furthermore, adequate sensory input is one key aspect of motor control and learning. As sensory input is known to be altered in children with CP, therapy approaches utilize appropriate sensory input, e.g. by facilitating task-specific physiological movement patterns, "placing" sensory-guided muscle activation to improve motor output (Levin et al. 2009; Mayston 2001; Trojan and Pokorný 1999). Although therapists apply these therapeutic methods regularly, scientific studies regarding the effectiveness of these methods to improve SVMC are scarce. The body of evidence is the largest for the adult stroke population and the effectiveness of upper extremity training in the paediatric population (Gordon 2016). The numbers of studies which investigated therapy-induced changes of SVMC in the lower extremity in children with CP is relatively limited (Chen et al. 2016; Pool et al. 2014; Karabay et al. 2015; Sukal-Moulton et al. 2014; Jung et al. 2013; Carraro et al. 2014; Degelaen et al. 2013; Burdea et al. 2013; Prosser et al. 2012; Bandholm et al. 2012; Cioi et al. 2011; Byanton et al. 2006; Buckon et al. 2002). Table 1 gives an overview of these studies. In none of these studies, SVMC was the main/primary outcome. The interventions focused mainly on improving strength and range of motion around the ankle joint (to

improve gait function). As these improvements in ankle muscle activity are closely related to motor control output, the authors were interested in monitoring SVMC as well. Thereby, the selected interventions either assessed changes in SVMC via indirect treatment approaches (medical or surgical reduction of spasticity)(Carraro et al. 2014; Degelaen et al. 2013; Bandholm et al. 2012; Buckon et al. 2002), direct approaches by means of progressive strength/coordination training (Jung et al. 2013); robot-assisted devices combined with virtual reality feedback (Chen et al. 2016; Sukal-Moulton, Clancy, L. Q. Zhang, et al. 2014; Meyer-Heim and Van Hedel 2014; Burdea et al. 2013; Burdea et al. 2011; Byanton et al. 2006) or functional electrical stimulation (FES) (Karabay et al. 2015; Pool et al. 2014; Prosser et al. 2012). Interestingly, no study investigating the effect of a traditional neurodevelopmental treatment approach such as Bobath on SVMC was identified.

However, in relation to the diversity of the interventions and measurement techniques of these studies as well as due to common methodological limitations (e.g. small and heterogeneous sample, lack of a control group, lack of previously evaluated outcome measure), the generalizability of these study results is limited. Nevertheless, these results show the trend of direct motor shaping approaches, which use sensory and visual feedback for lower limb.

Table 1: Interventions aiming to improve SVMC of the lower extremity in children with CP

Author	Intervention		Measurement	Effect	
	participants	duration	SVMC outcome measures	reported values	sig. change
Interventions aiming to improve SVMC <u>indirectly</u> – via decreasing spasticity					
Selective Dorsal Rhizotomy (SDR)					
Buckon et al. 2002	25 CP	6 and 12 months post SDR	sEMG plus isometric strength measures: co-contraction in the lower extremity	co-contraction during knee extension normalized post SDR, but not during ankle movement	positive trend
Carraro et al. 2014	9 CP	pre - post 12 months	Total SCALE score	pre SDR: 4(2-8) post SDR: 7(3-10)	yes
Botulinum Toxin Injection (BTX)					
Degelaen et al. 2013	14 CP bilateral 14 CP unilateral GMFCS I-II	4 months post BTX	interjoint coordination of hip and knee during swing phase, minimal relative phase analysis	only plots	stated as sign. but values not reported
Bandholm et al. 2012	14 CP (7 intervention vs 7 control) GMFCS I	12 weeks, 2 x training per week intervention group: custom-made equipment that allowed isolated ankle DE control group: normal/regular training	submaximal torque steadiness of isometric ankle dorsi- and plantarflexion	torque steadiness of DE improved sign, similarly in the two groups, plus reduction in antagonist (soleus) co-activity	yes

Abbreviations: SVMC: Selective Voluntary Motor Control; CP: Cerebral Palsy; SDR: Selective Dorsal Rhizotomy; BTX: Botulinum Toxin Injection; m. tib.ant: tibialis anterior muscle; m. gastroc: gastrocnemius muscle; sEMG: surface ElectroMyoGraphy

(Table 1 continued)

Author	Intervention		Measurement	Effect	
	participants	duration	SVMC outcome measure	reported values	sig. change
Interventions aiming to improve SVMC directly – via working on physiological muscle activation					
progressive strength training m. gastrocnemius					
Jung et al. 2013	6 CP	6 weeks, 3 x training per week	SMC and sEMG of m. tib.ant; m. gatsroc.	not reported	no
Neuromuscular Electrical Stimulation (NMES) and conventional physiotherapy for ankle muscles					
Karabay et al. 2015	28 CP GMFCS I-V	30 training sessions	SMC of the ankle joint only	pre: 3 (1-4) post: 3 (1-4)	no
Functional Electrical Stimulation (FES) device: “Walk Aide” on active ankle dorsiflexion					
Pool et al. 2015	32CP (unilat) GMFCS I-II	16 weeks, 3.3h (mean) a day	SMC of the ankle joint only SCALE more affected leg score	SMC Pre: 0.11 (0.09) Post: 0.20 (0.12) SCALE: Pre: 4.94 (1.12) Post: 6.06 (1.53)	sign. improvement in the intervention group, no in control group
Pool et al. 2014	12 CP (unilat) GMFCS I-II	16 weeks, 3.3h (mean) a day	SMC of the ankle joint only	pre: 2 (0-4) Post: 3.5 (1-4)	yes
Prosser et al. 2012	19 CP GMFCS I-II	4 months, 5.6h a day	3D gait analysis: kinematic mean and peak DE during swing phase (control group no effect)	pre: 1.9 (5.3) post: 3.1 (5.1)	yes
Virtual Reality augmented ankle robotic training device					
Chen et al. 2016	41 CP GMFCS I-III	6 weeks, 3 x training (laboratory and home-based)	SCALE leg scores	pre : 5.8 ± 2.7 post : 7.3± 2.1 follow-up : 7.3± 2.2	yes
Sukal-Moulton et al. 2014	28 CP GMFCS I-III	6 weeks, 2 (75 min) x training	SCALE leg scores	not reported	stated as sign. but values not reported
Cioi et al. 2011 and Burda et al. 2013	1 CP	12 weeks, 3 x training	DE / PF Torque and 3D ankle kinematics during gait	in-phase (s): pre: - 0.69; post: 0.10 Anti-Phase (s): pre – post: 0.76 0.42	yes
Wu et al. 2010 Wu et al. 2011	12 GMFCS I-III	6 weeks, 3 x training	SCALE (all joints) sEMG (m. tib. ant; m. gastroc.) Electrogoniometer: ankle ROM	not reported	stated as sign. but values not reported
Byanton et al. 2006	10 CP GMFCS 1-II	one 90 min training sessions	Electrogoniometer: active ankle DE ROM	greater mean ankle active ranges of motion into dorsiflexion during VR	yes

Abbreviations: SVMC: Selective Voluntary Motor Control; CP: Cerebral Palsy; SDR: Selective Dorsal Rhizotomy; BTX: Botulinum Toxin Injection; m. tib.ant: tibialis anterior muscle; m. gastroc.: gastrocnemius muscle; VR: Virtual Reality; NMES: NeuroMuscular Electrical Stimulation; FES: Functional Electrical Stimulation; sEMG: surface ElectroMyoGraphy; 2D: two dimensional; 3D: three dimensional; ROM: Range of Motion; DE: Dorsiflexion; PF: Plantarflexion; SMC: Selective Motor Control test; SCALE: Selective Control Assessment of the Lower Extremity; sign: significant

2.2.9 Neuronal structures and mechanisms involved in improving impaired SVMC

This section will describe the underlying neurophysiological mechanism, concepts and neuroanatomical structures which may be involved when training SVMC. Section 2.2.9.1 describes the concept and effectiveness of neuroplasticity. Section 2.2.9.2 summarises the two therapy principals that are based on the concept of neuroplasticity: constrained movement therapy and mirror therapy. How these therapies have revealed to show potential to improve SVMC in the upper extremity, even post UMN lesions, will also be discussed in this section.

2.2.9.1 Concept of neuroplasticity

As mentioned above, neuroplasticity is the major concept supporting recovery or (re-)training of selective voluntary movements in patients during neuro-rehabilitation (Nudo 2013). Neuroplasticity describes the “ability of the brain to form and reorganize synaptic connections, especially in response to learning or experience or following injury” (Nudo 2013). Hubel and Wiesel were the first scientists who discovered from their experiments on the visual cortex of kittens and monkeys that the development of the CNS is activity dependent and not predetermined, meaning that it will be shaped by the interaction with external factors (Forssberg 2014). The two commonly used slogans “Use it or lose it” and “neurons that fire together will wire together” nicely reflect this principle. This activity-induced plasticity is supported by developmental changes of the brain (e.g. disappearance of the ipsilateral CST projections) as well as by studies investigating cortical changes after training specific motor tasks. For instance, training the sensitivity of the fingers of monkeys will reorganise their cortical map in accordance with their trained fingers. Similarly, an increase of the cortical presentation of the joint needed for a repetitively performed activity has been observed in studies comparing professional athletes (e.g. badminton players) and musicians (e.g. pianists) with unprofessional controls. In neuro-(paediatric) rehabilitation, the concept of movement therapy has been mainly studied in unilateral stroke or CP involving upper extremity training like constraint-induced therapy or mirror therapy.

2.2.9.2 Neurophysiology of constraint-induced therapy and mirror therapy

Constraint-induced therapy involves constraining movements of the less affected arm, usually with a sling or splint, while intensively inducing the use of the more affected arm. Several studies have shown that constraint-induced therapy is associated with an use-dependent cortical reorganization that increases the area of cortex involved in the innervation of movement of the more affected limb (Johnston 2009).

Mirror therapy is another type of therapy that works on the principles of neuroplasticity and also makes use of another neural system (mirror neuron system) (Small et al. 2012; Buccino et al. 2006; Bhattacharya and Lahiri 2002). In mirror therapy, the affected limb is placed behind a mirror, which is placed in such a way, that the reflection of the opposing (unaffected) limb appears in place of the hidden limb (Buccino et al. 2006). The patient is then asked to perform functional movements with the unaffected side while simultaneously looking into the mirror. As the brain prioritises visual feedback over somatosensory feedback concerning limb position, it receives the illusion that the affected limb is moving physiologically. Thereby, motor processes in the brain will become stimulated in the same manner as when the patient would have moved the limb. The mirror neuron system of the brain is the primary cause of this action. Although the understanding of this complex system is still in its infancy, the mirror neuron system is known to play an important role during observational learning and becomes activated in the brain immediately after birth (Buccino et al. 2006; Bhattacharya and Lahiri 2002). For instance, during the earliest stages of development, this system enables the baby to mimic gestures of its parents (Buccino et al. 2006). In later stages, it supports learning of more complex motor (e.g. talking, walking) and social behaviours, also by activating CPGs (Forssberg 2014). During mirror therapy, mirror neurons will become activated as this system is also involved with the process of laterality reconstruction (e.g. ability to differentiate between the left and the right side). Although mirror neurons are activated by visual input, research has shown that this system fails to be activated when the observer has no interest in the shown movement task or in those with autism. Mirror therapy has shown to be effective in conditions such as phantom limb pain, chronic pain syndromes and stroke and seems to normalize the neuropathic processes which cause issues with pain and disturbed movement control (Bhattacharya and Lahiri 2002; Sütbeyaz et al. 2007).

As children with CP, in contrast to adult stroke patients, cannot rely on physiologically movement experience of their affected limb, the effectiveness of mirror therapy in this population needs to be investigated (Forssberg 2014). A recent fMRI study on children with unilateral spastic CP showed that solely by observing the movements of their paretic hand, similar brain areas of the mirror neuron system (bilateral inferior and superior parietal region) were activated in the same way as in neurologically intact humans (Dinomais et al. 2013). There are only a few studies on the use of mirror therapy for the lower extremity and all included adult stroke survivors (Sütbeyaz et al. 2007; Small et al. 2012). Intervention studies on impaired upper extremity control in children with CP have shown that for instance home-based daily mirror therapy training resulted in significant improvements in strength and function of the paretic upper limb. Nevertheless, these improvements were not significantly better compared to the conventional home-training (Bruchez et al. 2016; Park et al. 2016).

The RST has shown to play a role in the restoration and rehabilitation of impaired motor control (Cahill-Rowley and Rose 2014; Hicks and Onodera 2012). Ontogenetically, the RST is an older brain structure. Recent studies in animals and humans suggest that the RST is able to partially compensate for the loss of CST control. Studies with dogs and monkeys have demonstrated the recovery of motor function disabilities from CST lesions whereby the RST was spared. If the RST was injured in addition to a lesion of the CST, motor deficits would be permanent (Forssberg 2014). Furthermore, diffusion tensor imaging (DTI) studies in patients with stroke and a lesion of the CST showed higher fractional anisotropy of the red nucleus in the affected hemisphere during the acute and chronic phase of recovery. Further, the bilateral red nuclei were the only regions with greater fractional anisotropy during the chronic phase when compared to the uninjured control group (Hinkley et al. 2009; Yao et al. 2009). Functionally, an increased red nucleus fractional anisotropy correlated with better motor function in the patient group (Hinkley et al. 2009). These results suggest that increased RST activity compensates for CST damage in stroke patients. What role the RST could play in children with CP for compensating impaired motor function has not yet been investigated. From the perspective of brain development, it is suggested that with the increased maturation of the CST, infant movements become less synergistic, and thus rely less on the RST (Cahill-Rowley and Rose 2014).

Overall, the same concepts of neuroplasticity and motor learning fundamentally exist, whether or whether not a UMN lesion is present. The underlying neural structures and networks (motor cortices, mirror neurons, and RST) might be altered, but through adequate input these are also shape-able and thus allow for restoration and rehabilitation of impaired motor control. Therapeutically, this implies that selective movement training should take place in a functional and attractive and stimulating context for the patient (Meyer-Heim and Van Hedel 2014; Lotze 2003; Trojan and Pokorný 1998). This indicates that therapies such as mirror and constrained induced therapy, which involves functional training as well as additional visual and sensory stimuli (for guiding motor learning), could potentially promote neuroplasticity and motor learning. Furthermore, synaptic consolidation of learning is only possible after “enough” repetitions (“Neurons that fire together, wire together.”) have been performed. In summary, everybody can learn to play piano with the feet, as long as enough motivation, time (repetitions) and feedback is given.

2.3 Outcome measures

Measurements in medical science and clinic allow to quantify the physical or behavioural characteristics of patients (Wright and Majnemer 2014; De Vet et al. 2011). The quantification of an impairment enables the researcher or clinician to collect specific, pre-defined information (data). The accumulation of this data supports the description and understanding of the measured characteristics outcome within the population of interest. Thereby it can help to understand a certain impairment (e.g. symptom) better or to evaluate the patient’s response (i.e. outcome) to a particular intervention. Outcome measures are defined as “tools that may be used to assess changes in particular attributes that are meaningful to a person’s life over time” (Ferguson 2017). The term ‘outcome’ reflects that this tool measures the result of a determinant (e.g. treatment, program, service) over a certain time phase. When applied before and after an intervention, it measures if and how much change occurred in the outcome variable. This knowledge is important for clients, service providers, clinical managers, policymakers as well as for researchers (Wright and Majnemer 2014). In order to select the “right tool” for any planned investigation, one should define the following testing criteria: i) purpose of the measurement (e.g. hypotheses: diagnostic, evaluative, prognostic), ii) population of interest (e.g. for which patient group was the tool originally developed?), iii) quality of the psychometric

properties, iv) intervention purpose (dependent variable), iv) feasibility of the measurement (e.g. costs, time) (Wright and Majnemer 2014; De Vet et al. 2011).

Measurement has been defined as the “*process of assigning numerals to variables to represents quantities of characteristic according to certain rules*” (Portney and Watkins 2000). The variable of interest can either be continuous (the variable can take any value along a continuum of a defined range), discrete (the variable is measured in units / whole numbers) or dichotomous (qualitative variables, which can only take two values). When the variable can be observed during the measurement process (e.g. the range of motion), it is known to be of direct nature. As most characteristics are not directly observable during the measurement (e.g. the heartbeat, muscle strength, SVMC), they have to be measured indirectly. Furthermore, a measurement is defined by its scaling system, which applies certain rules/values to the variable of interest. The scale of measurement can either be nominal, ordinal, interval or ratio (Portney and Watkins 2000). The lowest level of measurement is the nominal scale (e.g. male = 0 or female = 1), which allows to classify people according to some criterion. Ordinal scales organise data into adjacent categories (e.g. SVMC is either normal = 2 or impaired = 1 or unable = 0) according to a severity-range. The main limitation of this scaling system is that its ordinal values do not represent quantity, meaning that they lack arithmetic properties (Portney and Watkins 2000; Bishop and Herron 2015). The next level of scaling is the interval scale wherein the intervals between the units of the measurement are known and of equal distance but are not related to a true zero (e.g., temperature in degrees Celsius). Only when using a ratio scale, a score in which zero represents total absence of the measured value, it will allow for a comprehensive interpretation. Therefore, classical guidelines recommend to apply mathematical and statistical operations only to interval or ratio scaled data and to stick to frequency counts for ordinal and nominal measures (Portney and Watkins 2000). However, there are a lot of examples in clinical and behavioural science in which ordinal data, for example data from validated questionnaires, has been used for statistical analysis. Ordinal measures are frequently used in medical and behavioural sciences for the reasons of clinical feasibility regarding time, costs, and expertise; some variables cannot be measured on another scale. Advocates for handling ordinal data arithmetically (as if they were interval data) suggest that a statistical procedure should be applied according to what is meaningful data and not strictly by the scale used (Norman 2010; Velleman et al. 1993). They argue that statistical analyses are not an end in themselves, but rather a

means to an end, enabling investigators to think about the data. Despite this ongoing statistical debate, the decision on the best outcome measures the researchers' hypotheses that the appropriate statistics depend on (De Vet et al. 2011). On behalf of a robust interpretation of the study results, the researcher has to justify the chosen outcome measures and statistical procedure used.

2.3.1 Psychometric properties

Psychometric properties are defined as “quantifiable attributes (e.g. validity, reliability) that relate to the statistical strength or weakness of a test or measurement” (De Vet et al. 2011). Knowledge about the quality of the psychometric properties of outcome measure is essential for clinicians and scientists to estimate the robustness and correctness of the measurement (De Vet et al. 2011; Wright and Majnemer 2014). Only by knowing that the chosen measure is measuring what it is intended to measure (validity) and by knowing its error-range concerning its reproducibility (reliability), the tester can estimate whether a measured change is a real change. Reliability and validity are interlinked since reliability is a prerequisite for validity (Portney and Watkins 2000). Another important property, especially regarding outcome measures, is responsiveness. Responsiveness evaluates the measure's sensitivity to change (Portney and Watkins 2000).

The COSMIN checklist (Mokkink and Terwee 2010) was developed to provide standard guidelines for evaluating the methodological quality of any clinically or scientifically applied measure. The COSMIN group used a Delphi study to gain agreement on how psychometric properties are defined and how their quality should be evaluated (Mokkink and Terwee 2010). Although the checklist was originally developed for studies on the measurement properties of health-related patient-reported outcomes, its application is also relevant for studies evaluating psychometric properties of other health measurement instruments (Mokkink and Terwee 2010). The main goal of the COSMIN group is to enhance the quality of studies being published and thereby improving the level of evidence for clinicians and researchers. The definitions of psychometric properties and instructions on how these should be assessed are of paramount importance to the main aim of this thesis: to identify a reliable, valid, and sensitive outcome measure of SVMC. Therefore, the following sections will discuss each psychometric property and its quality criteria. A summary of the definitions proposed by the COSMIN group is listed in table 2.

Table 2: Domains and definitions of measurement properties

<i>Domain</i>	<i>Measurement property (COSMIN Box)</i>	<i>Definition</i>
<i>Reliability</i>		The degree to which the measurement is free from measurement error: The extent to which scores from patients who have not changed are the same for repeated measurements under several conditions, e.g., using different sets of items from the same health related-patient reported outcomes (i.e. outcome measure) (internal consistency), over time (test re-test) by different persons on the same occasion (inter-rater) or by the same persons (i.e. raters or responders) on different occasions (intra-rater)
	Reliability (Box A)	The proportion of the total variance in the measurements which is because of “true” differences among patients
	Internal Consistency (Box B)	The degree of the interrelatedness among the items.
	Measurement error (Box C)	The systematic and random error of a patient’s score that is not attributed to true changes in the construct to be measured.
<i>Validity</i>		The degree to which an instrument measures the construct(s), it purports to measure.
	Content (Face) (Box D)	The degree to which the content of an instrument is an adequate reflection of the construct to be measured. (Degree to which (the items of) an instrument indeed look(s) as though they is/are an adequate reflection of the construct to be measured.)
	Construct (Structural) (Box E-G)	The degree to which the scores of an instrument are consistent with hypotheses (for instance with regard to internal relationships, relationships to scores of other instruments, or differences between relevant groups) based on the assumption that the instrument validly measures the construct to be measured. (Degree to which the scores of an instrument are an adequate reflection of the dimensionality of the construct to be measured.)
	Criterion (Gold Standard) (Box H)	The degree to which the scores of an instrument are an adequate reflection of a “gold standard”
<i>Responsiveness</i>	Responsiveness (Box I)	The ability of an instrument to detect change over time in the construct to be measured

2.3.2 Measuring the quality of studies presenting psychometric data

The COSMIN checklist has shown to be a suitable tool to evaluate the methodological quality of studies evaluating the psychometric properties in health-related patient-reported outcomes (e.g. questionnaires for assessing symptoms, functional status, health-related quality of life). Additionally, it is suitable for objective outcome measures within the neuro-paediatric field (Gerber et al. 2016; Ammann-Reiffer et al. 2014). The checklist consists of the three domains: reliability, validity, and responsiveness. Each domain contains one or more measurement properties (COSMIN box). The evaluation of each property is performed by scoring general (e.g. handling of missing values, sample size) and property-specific items given within each

COSMIN box (Appendix 1). A 4-point rating scale ('excellent', 'good', 'fair', 'poor' or 'not applicable') is used for scoring each item. The lowest score of all items of chosen particular 'COSMIN box' (e.g. for reliability) determines the overall methodological quality of the psychometric property. In addition to the rating of the methodological quality of a study assessing one or more measurement properties, Terwee et al. (2007, 2010) proposed to rate the quality of the measurement properties themselves. This proposed rating scale is shown in Table 3.

In the following sections of this thesis, both the COSMIN checklist and the Terwee quality criteria for each domain will be described in more detail.

Table 3: Quality criteria for measurement properties

Property	Rating	Quality criteria
Content validity (Box D)	+	A clear description is provided of the measurement aim, the target population, the concepts that are being measured, and the item selection AND target population and (investigators OR experts) were involved in item selection
	?	A clear description of above-mentioned aspects is lacking OR only target population involved OR doubtful design or method
	-	No target population involvement
Criterion validity (Box H)	+	Convincing arguments that gold standard is "gold" AND correlation with gold standard >0.70
	?	No convincing arguments that gold standard is "gold" OR doubtful design or method
	-	Correlation with gold standard <0.70, despite adequate design and method
Construct validity (Box E,F,G)	+	Specific hypotheses were formulated AND at least 75% of the results are in accordance with these hypotheses
	?	Doubtful design or method (e.g. no hypotheses)
	-	Less than 75% of hypotheses were confirmed, despite adequate design and methods
Reliability (Box B)	+	ICC or weighted Kappa>0.70;
	?	Doubtful design or method (e.g. time interval not mentioned)
	-	ICC or weighted Kappa <0.70, despite adequate design and method
Responsiveness (Box I)	+	Correlation with an instrument measuring the same construct >0.50; OR at least 75% of the results are in accordance with the hypotheses OR AUC>0.70 AND correlations with related construct is higher than with unrelated constructs
	?	Solely correlation determined with unrelated constructs
	-	Correlation an instrument measuring the same construct <0.50; OR <75% of the results are in accordance with the hypotheses; OR AUC<0.70 AND correlations with related construct is higher than with unrelated constructs

Abbreviations: += positive rating; ?=indeterminate rating; -=negative rating; ICC= Intraclass correlation coefficient, AUC= area under the curve of the receiver operating characteristics curve; adapted from Terwee et al. (2007)

2.3.3 Reliability

Reliability is defined as the degree to which the measurement is free from measurement error (Portney and Watkins 2000). There are three subtypes of reliability addressing slightly different aspects of this psychometric property: reliability, internal consistency and measurement error (De Vet et al. 2011). Internal consistency or homogeneity define as the degree of the interrelatedness among items. As this type of reliability is only relevant for the evaluation of questionnaires, it will be discussed in this thesis no further.

Reliability concerns the reproducibility of the results, i.e. when a test is repeated over time on the same participants (test-retest reliability), when used by different assessors on the same occasion (inter-rater reliability), or by the same assessor(s) on different occasions (intra-rater reliability). When the agreement is high within the sample, the variance is low, meaning that the instrument is measuring precisely. Thereby reliability represents how much of the total sample variance is attributable to true differences between scores; it can be expressed by the following coefficient (Portney and Watkins 2000):

$$\frac{\text{true score variance}}{\text{true score variance} + \text{error variance}}$$

Reliability is excellent, when the coefficient is 1.00, meaning that the error is zero. As the errors increase, the ratio approaches zero and the reliability of the tool diminishes. Statistically, reliability can be expressed by measures of correlation and/or agreement. While correlations reflect the degree of how scores vary together (i.e. consistency), they are not a measure of the extent of agreement between the two sets of measurements. To establish that repeated tests result in the same values, estimates of agreement are needed. The COSMIN checklist for reliability recommends to apply an Intraclass Correlation Coefficient (ICC) (ICC model 2: inter-rater; ICC model 3: intra-rater) for all continuous scores and to use a (weighted) Kappa for all other types of scores (Mokkink and Terwee 2010).

According to the quality criteria from Terwee et al. (2007), the reliability of a certain measure is rated *positive*, when the ICC/or weighted Kappa is above 0.7. An *indeterminate* rate is given, *when* the study assessing reliability was methodologically flawed (COSMIN rating of 'poor'), allowing bias to occur (e.g. time interval between

the measures was not stated or selected appropriately, participants condition was not stable between the two measurement occasions) (Appendix 1: Box B). If the design and method were adequately chosen, but the ICC or weighted Kappa was below 0.70, reliability would be rated *negative* (Terwee et al. 2007).

When the agreement between the true value and the observed value is low, the source of “noise” that gets in the way of finding the true score must be detected. **Measurement error(s)** can occur either systematically (e.g. false calibration of the instrument) or randomly (e.g. fatigue). While systematic errors can be corrected, for instance by recalibrating the system or by adjusting the over- or underestimated measured value, the handling of random errors is more challenging, since these occur due to chance (e.g. mistakes, inattention) (Portney and Watkins 2000). The general assumption with random errors is that they will eventually cancel each other out, if enough measurements are taken. To minimize both types of error, the methodological testing procedure should be standardized, taking the common sources of errors into account. These errors include errors within the measurement (e.g. for technical measures: calibration error); rater/tester (e.g. different levels of work experience); subjective differences within the sample; environmental differences (Portney and Watkins 2000) such as temperature or time or the time interval between the repeated measures. De Vet al. (2011) recommends the following analytic parameters to represent measurement error:

- Standard Error of Measurement (SEM) = $SEM = SD \sqrt{(1-r_{xx})}$: measures test scores that are spread around a “true” score
- Smallest Detectable Change (SDC) = $1.96 \times \sqrt{2} \times SEM$: measures the variation in a scale due to measurement error
- Limits of Agreement (LoA) = (95%CI of the mean difference $\pm 1.96 \times SD$ of the differences): represents the 95% confidence interval for both the upper and lower limits of agreement (Bland–Altman method).

When evaluating the methodological quality of a study investigating the measurement error of an instrument in accordance to quality criteria of Terwee et al. (2007), similar items are applied, with an exception for the statistic items (Appendix 1: Box C).

2.3.4 Validity

The validity of a measurement tool describes to which extent an instrument measures what it is intended to measure (Portney and Watkins 2000). The COSMIN checklist includes five different boxes within the validity domain: content validity, structural validity, hypothesis testing (construct validity), cross-cultural validity and criterion validity (Mokkink and Terwee 2010).

The first form of validity is **face validity**, which reflects the “degree to which a measurement instrument looks as though it is an adequate reflection of the construct to be measured” (Mokkink and Terwee 2010). As this form of validity is dependent on subjective assessment, a quantified evaluation is not possible. Nevertheless, it is of great importance since an instrument or questionnaire, which seems to have low face validity, should not be further validated. Therefore, it is not separately listed within the COSMIN checklist but together with **content validity**. This is the next higher form of validity and uses expert-panel feedback as well as previously validated frameworks such as the ICF-CY to assess whether the measurement instrument (e.g. items, scaling system) adequately represents the construct it is aiming to represent. Evaluation of this type of validity focuses on how concrete the construct of interest and the chosen items are defined. Furthermore, the purpose of the instrument and the population of interest should be described in detail (De Vet et al. 2011) (Appendix 1: Box D).

After the face and content validity of an instrument have been confirmed, the validation process should continue by testing either its **criterion validity** or **construct validity**. If a gold standard measure is available, criterion validity should be evaluated. Criterion validity is defined by the COSMIN panel as “the degree to which the scores of a measurement instrument are an adequate reflection of a gold standard”. It is further subdivided into concurrent validity (the assumption that the test scores and gold standard scores are the same) and predictive validity (the assumption that the test scores predict the scores of the gold standard). The recommended statistical procedure depends on the measurement levels of the gold standard and the measurement instrument tested (Appendix 1: Box H) (De Vet et al. 2011). As a gold standard is often lacking in the field of paediatric neurorehabilitation, construct validity should be used to estimate the validity of the instrument. For the assessment of any type of construct validity (structural validity, hypotheses testing, cross-cultural validity), a priori formulated hypotheses about the (expected) relationships between

the instrument scores and another instruments or groups are tested (De Vet et al. 2011; Mokkink and Terwee 2010; Portney and Watkins 2000).

Structural validity investigates how and if the scoring system of a measurement tool is accurate through factor analysis (classical test theory or item response theory) (De Vet et al. 2011) (Appendix 1: Box E).

When **hypotheses testing** is performed to evaluate the validity, a priori formulated hypotheses about the relationship of the measurement scores and the scores of comparator instrument or group are tested. The expected relationship could either be that both tools measure the similar construct (*convergent validity*) or that the instruments are measuring different constructs (*discriminant validity*). The robustness of this type of validity is strongly dependent on the previously defined construct of interest and the formulated hypotheses (Appendix 1: Box F).

Cross-cultural validity has to be assessed when a measurement tool is translated into another language or has been adapted due to cultural differences. It tests the “degree to which the performance of the items in a translated or culturally adapted instrument is an adequate reflection of the performance of items in the original version of the instrument” (Mokkink and Terwee 2010). The COSMIN panel recommends the following translation process for measurement instruments (Beaton et al. 2007): i) translation of the instrument by two independent native translators who have the target language as their mother tongue; ii) creation of consensus version; iii) back-translation into the original language by two other translators with the original language as their mother tongue and who are blinded to the original version, (iv) obtain approval by an expert committee meeting of all translators and scientists and clinicians involved; v) pilot testing of a pre-final version within a small sample; vi) obtain approval by the developers of the original version. Following this translation process, validation of the translated or culturally adapted instrument should be approved by testing expected correlations with related tools. For instance, it should be assessed whether a given measure is interpreted in a conceptually similar manner by respondents representing different genders or different cultural backgrounds. Statistically, this is known as the assessment of *measurement invariance* and commonly tested by factor analysis, logistic regression and item response theory techniques (Appendix 1: Box G).

In accordance to the quality property criteria from Terwee et al. (2007), all of the three construct validity forms are rated positive, if specific hypotheses have been formulated and if at least 75% of the results are in accordance with these hypotheses. If no hypotheses were formulated, the design will be doubtful, or the construct validity will be rated as 'indeterminate'. A 'negative' rating is given if the design was adequate, but less than 75% of hypotheses were confirmed.

2.3.5 Responsiveness

While validity refers to the validity of a single score, responsiveness refers to the validity of a change score. Therefore, COSMIN considers responsiveness as an aspect of validity, which evaluates "the ability of an instrument to detect change over time in the construct to be measured" (Mokkink et al. 2010). Accordingly, whenever a scientist or a clinician aims to detect a change in a patient's health condition due to recovery or intervention, responsive outcome measures should be used. Depending on the availability of a gold standard, responsiveness is tested either via criterion validation or construct validation. The main hypothesis within responsiveness testing is that the changes in the patients' health status regarding the construct of interest are in the expected direction and magnitude in both instruments (new instrument versus gold standard/comparator instrument) (De Vet et al. 2011). Due to the focus on changes scores, a longitudinal study is required in which at least some of the patients are known to change on the construct of interest (e.g. due to natural recovery or intervention) and with at least two measurement occasions. Methodologically, the time interval between the two measurements as well as occurring events (e.g., interventions) within this time interval need to be stated and described. If a particular intervention is selected to assess responsiveness, the evidence regarding its effectiveness should have been established in advance. If an adequate gold standard is available, the a priori hypotheses should define the expected level of agreement between changes in the outcome measure of interest and changes on the criterion. Additionally, the strength of the relationship between the changes in scores of both tools should be defined. When construct validation is performed, the selected comparator measurement tool or group should be described in detail, including its psychometric properties of the comparator instrument (Mokkink and Terwee 2010). The statistical analysis of responsiveness depends on the predefined hypotheses and testing design (Appendix 1: Box I). Effect sizes, paired t-test and Guyatt's responsiveness ratio are considered inappropriate to assess responsiveness,

because these parameters test either the magnitude or statistical significance of the change score or refer to the interpretability of the change score, but do not test its validity (De Vet et al. 2011).

The quality of the responsiveness of a certain measure is rated positive if the correlation coefficient with an instrument measuring the same construct is above 0.50; if at least 75% of the results are in accordance with the hypotheses, or if the area under the curve of the receiver operating characteristics curve is above 0.70 and the correlations with related construct are higher than with unrelated constructs. Consequently, responsiveness quality is rated negative, if the values are below the previously given thresholds. If solely correlation coefficients are determined with unrelated constructs, responsiveness will be rated as 'indeterminate' (Terwee et al. 2007).

2.3.6 Outcome measures for measuring SVMC

Section 2.2.8 discussed the lack of scientific evidence on nature as well as the trainability of SVMC (Cahill-Rowley and Rose 2014; Dobson 2010). A possible explanation for this gap in knowledge might be related to the problems in measuring SVMC (Gordon 2016; Fowler 2010; Dobson 2010; Zwaan et al. 2012). One problem is the interrelated pathological nature of UMN signs (Cahill-Rowley and Rose 2014). Section 2.2.7 discussed the interrelated nature of SVMC with other UMN signs as well as with other impairments of body function (e.g. lack of trunk control, balance problems). Due to the interrelated nature of UMN signs, an isolated investigation of SVMC is quite challenging in the practical measurement context. Questions like "How to differentiate between a lack of SVMC and muscle strength or spasticity?" must be taken into consideration to undertake a precise measurement (Fowler et al. 2009). One way to control for interfering effects of other impairments, for instance muscle weakness or spasticity, is to measure these two impairments independently. Another important aspect is that the tool itself is developed in a way that it controls for the most interfering factors (Portney and Watkins 2000).

Another problem when measuring SVMC is that although the movement can be observed, the underlying controlling structure and processes are hidden. The ability to selectively activate a muscle or muscle group within a given movement or posture determines how the movement is performed. As most of our activities of daily life are complex, three-dimensional movements, which take place simultaneously over

multiple joints, measuring *how* the movement is performed is more challenging than measuring the quantity. For example, it is quite easy to assess, if and how fast a patient can stand up, but it is more complex to measure *how* he/she is standing up from sitting. For instance, a child could stand up physiologically by slightly leaning forward and lifting up by progressively straightening knees and hips; or in an unphysiological manner by e.g. leaning the trunk backward to extend the hips and leg and then pulling up with the help of the hands. The majority of available assessment tools are on the activity level and fail to measure how the patient is performing a motor task per se (Levin et al. 2011).

Concerning the lack of SVMC outcome measures, the next two sections will summarise the currently available outcome measures of SVMC in children with CP, as well as in the adult population of patients with UMN lesion.

2.3.6.1 Measures of SVMC in children with CP

Literature revealed that in 1999 Boyd and Graham's Selective Motor Control (SMC) test was the first test that aimed to measure SVMC. The authors were interested in measuring the changes in selective ankle control after Botulinum toxin injections (Boyd and Graham 1999). Later, Trost modified this assessment and also included the assessment of the selectivity around the knee and hip joints (Smits et al. 2010; Löwing and Carlberg 2009). In 2009, Fowler and colleagues developed the SCALE for assessing SVMC of the hip, knee, ankle, subtalar joint and toes (Fowler et al. 2009). Although the SCALE and the SMC have similarities, as they both assess selectivity around the lower limb joints, the scoring and test procedure, as well as how they handled muscle weakness, differs. The SCALE does not only assess selective ankle movement but also subtalar and toe movement. Furthermore, its scoring system and testing position allow to differentiate between muscle weakness and selectivity (Fowler et al. 2009). The patient is asked to move the joint on verbal instruction, which allows to assess voluntary motor control. These assessment characteristics indicate high content validity of the SCALE concerning the definition of SVMC. This clinical assumption was confirmed by a mean content agreement of 91.9% from 14 experienced clinicians (Fowler et al. 2009). Construct validity of the SCALE was supported by a significant inverse correlation ($r = 0.83$, $p < 0.01$) with the participants' GMFCS level in a study including children with CP (Fowler et al. 2009). Furthermore, the SCALE showed that impairment of SVMC increases from proximal to distal in

children with CP (Fowler et al. 2010) and that its score strongly correlates with swing knee extension acceleration during gait ($r = 0.85$, $p < 0.001$) in the same patient population (Goldberg et al. 2011). Interrater reliability of 6 clinicians rating SVMC of 20 participants with the SCALE was high, with ICC ranging from 0.88 to 0.91 ($p < 0.001$) (Fowler et al. 2009). According to these study results and in comparison to available alternatives, the SCALE is currently the most common assessment tool of choice for testing SVMC in children with CP (Dobson 2010). Nevertheless, as its scoring system is originally designed for diagnostic and prognostic and not for evaluative purposes (i.e. effects of treatment), its clinical application is limited (Fowler et al. 2009). To improve SVMC in children with UMN lesion, health professionals need to evaluate the effects of their interventions, and for this measurement instruments with an evaluative character are needed (Zwaan et al. 2012; Dobson 2010).

2.3.6.2 Measures of SVMC in other populations with a UMN lesion

Because of the lack of appropriate SVMC outcome measure in the neuropediatric population, the literature searched for alternative tools within the entire population of patients with UMN lesion. The aim of this systematic review was to create an up-to-date overview of measures, which already have been reported and used to measure and/or evaluate SVMC of the lower extremity in patients with UMN lesions.

A MEDLINE search revealed 1572 hits (Appendix 2). After screening titles and abstracts, 112 papers were included. These papers reported on outcome measures of SVMC ranging from simple observational tools to complex procedures like Transcranial Magnetic Stimulation (TMS) or functional Magnetic Resonance Imaging fMRI (Appendix 3). Several observation-based tools were identified for the assessment of SVMC of the upper extremity (Wolf et al. 2001; Levin et al. 2004; Van de Winckel et al. 2006), the lower extremity and the overall body function (Collen et al. 1991; Gowland et al. 1993; Sødning et al. 1996; Chino et al. 1996; Ahemd et al. 2003; Liu et al. 2002; Miller et al. 2008) in patients with stroke. Many assessments, like the Fugl-Meyer motor assessment (Gladstone et al. 2002), include subsections that specifically test SVMC of the lower extremity. Their scoring is similar to the ordinal scoring form with the paediatric assessments (SCALE; SMC) ranging from “0” = not able to move at all, over “1” = for impaired selective movement pattern, to “2” = for normal selective movement. As these tests were designed mainly to evaluate the initial motor control recovery of patients with stroke, their scoring system might be

suitable (i.e. sensitive) for detecting improvements of SVMC, but the ordinal scoring likely lacks sensitivity to evaluate therapy effects on SVMC in children with congenital brain injuries (Fowler et al. 2009). Additionally, their testing procedure is designed for the assessment of patients with hemiparesis only and is therefore not suitable for patients with diplegic or tetraplegic presentation of symptoms (Fowler et al. 2009). Hence, because of the limitations discussed above, none of these stroke assessment tools are indicated for either discriminative or evaluative assessment of SVMC in children with UMN lesion.

Laboratory-based SVMC measures include tools which test any form of active range of motion (Pomeroy et al. 2003), force and torque (e.g. Dynamometer (Engsberg et al. 2001; Braendvik et al. 2013), muscular activity (e.g. surface Electromyography (sEMG): on-off ratio (Doornik et al. 2008)); latency and activation time; Frequency/Wavelet Analysis (e.g. Prosser et al. 2010; Lauer et al. 2007); voluntary response index (Lee et al. 2004), as well as a combination of muscular activity and torque tracking (i.e. Knutson et al. 2011; van Hedel et al. 2010; Kiyama et al. 2011). Considering that SVMC and muscle weakness are different but also related impairments, measurements aiming to detect minor degrees of muscle weakness (submaximal strength) or using submaximal-torque tracking exercises are measuring selective motor neuron activation within the muscle, and thereby SVMC (van Hedel et al. 2010). Other laboratory SVMC measures use more complex measurement procedures focusing on the ICF body structure level, such as neuroimaging measures used for motor map localization (e.g. Wittenberg 2009; Möller et al. 2005) and TMS measuring corticospinal connections (e.g. Kesar et al. 2012; McKay et al. 2005). These papers provide an interesting insight into the underlying basic pathology (e.g. TMS norm values for children with CP (Kesar et al. 2012) and intervention-induced changes in SVMC (Everaert et al. 2010). The psychometric properties of sEMG (Zwaan et al. 2012), three dimensional gait analysis (Fowler et al. 2010; Fowler and Goldberg 2009; Goldberg et al. 2011) and isometric torque measurement (Bandholm et al. 2008) have been assessed (validity, reliability) in children with CP as well.

Only the minority of measures had been tested on their psychometric properties: validity, reliability, and responsiveness. Especially for laboratory-based SVMC measures such as sEMG or torque measurements, evidence of their validity and responsiveness is lacking, despite the fact that they have been used in interventional studies.

Although the psychometric evidence for many instruments was limited, several measures seemed to be interesting, because they fulfil one or more criteria of a theoretically ideal SVMC measure: i) measuring SVMC in accordance to its definition (focusing on single-joint movement and muscle activation) (Sanger et al. 2006), ii) applicable also for bilateral involvement, iii) reference/norm values available, iv) child friendly application (e.g. short duration, non-invasive) and v) well reliable, valid and responsive. For example, sEMG measures allow to directly measure voluntary activation of a muscle. Even in patients with little muscle strength, sEMG can detect voluntary activations without actual or only small joint movements. sEMG analyses can also provide in-depth information on activation of the agonist, antagonist and synergists (e.g. on-off ratio (Downing et al. 2009; Olsen et al. 2013; Damiano et al. 2000); latency and activation time (Tedroff et al. 2006); frequency or wavelet analyses (Lauer et al. 2005; Lauer et al. 2007)) and the voluntary response index (Zoghi et al. 2013; McKay et al. 2005; Lee et al. 2004). Combining sEMG with kinematic measurements further allows assessing co- and mirror movements and the percentage of the movement cycle for which a muscle acts concentrically, eccentrically, isometrically or is inactive. This SVMC measuring approach was used by Perry et al. (2001), who investigated the effect of dorsal selective Rhizotomy on voluntary muscle activation in non-ambulatory children with CP during a cyclic movement with their leg while supine. When sEMG is combined with (submaximal) torque tracking (i.e. Damiano et al. 2000; van Hedel et al. 2010; Kiyama et al. 2011) or isometric force measures (e.g. dynamometry (Lebiedowska and Fisk 2003; Lebiedowska et al. 2004)) additional information about the patient's voluntarily ability of submaximal force regulation can be gained. Again, this also allows the assessment of SVMC in relatively weak patients. Using this measurement strategy, van Hedel et al. (2010) observed that SVMC of ankle dorsiflexion was mainly affected in patients with stroke, but remained largely unaffected in patients with incomplete spinal cord injury, despite both groups of patients experienced comparable muscle weakness (Van Hedel et al. 2010). Another interesting combination of techniques was used in the "Brain Motor Control Assessment" (McKay et al. 2005; Lim et al. 2005). The "Brain Motor Control Assessment" measures the motor output from the central nervous system during a variety of reflex and voluntary motor tasks performed under controlled conditions using sEMG and TMS. One of its main outcomes, the "Voluntary Response Index", is calculated from quantitative analysis of the comparison of sEMG data during defined voluntary movement in patients versus neurologically intact people (Lee et al.

2004). We consider this as one of the most promising SVMC measures because it has been validated and uses healthy reference values for normalization. Unfortunately, to our knowledge, the “Voluntary Response Index” has not been yet evaluated in children with UMN lesions.

To conclude, although some measures for SVMC of the lower extremity are available in the population of patients with UMN lesions and have also been used to evaluate therapy-induced changes of SVMC, their level of psychometric evidence has not yet been evaluated systematically. Since the assessment of SVMC in children with CP differs from that in adult patients with UMN lesions, interchanging measurement tools will not be possible, or only after several necessary adaptations are made. In people with CP, the SCALE seems to be a very promising assessment tool. Nevertheless, its scoring system likely lacks sensitivity to detect therapy-induced changes. Using the SCALE’s testing procedures while recording neurophysiological measurements (e.g. sEMG, isometric torque measurement or TMS) seems to be a possible way to evaluate therapy-induced changes in SVMC.

Nevertheless, to be able to choose the most valid, reliable, sensitive as well as practical tool for evaluating changes in SVMC in children with UMN lesions, a systematic review on psychometric properties of available tools is necessary.

2.4 Rationale for the overall aim(s) of the thesis

Summarizing the above stated main research findings and gaps in knowledge about SVMC of the lower extremity in children with CP, reveals the following key points:

- The pathophysiological nature of SVMC in relation to other UMN signs and their interrelation to gross motor function is complex and yet not fully understood.
- Recent research indicates that SVMC is an important predictor of gross motor function (including gait) in children with CP.
- There are only a few studies investigating the effects of therapy on SVMC, and there is some evidence of the efficacy of these interventions.
- Concepts of neuroplasticity and motor learning underpin the trainability of SVMC even in patients with UMN lesions.

- Both observational and laboratory-based measures for SVMC of the lower extremity are available for children and adults with UMN.
- Validity and reliability of the SCALE have been assessed but its ordinal scaling may be too broad to measure therapy-induced changes of SVMC.
- Currently, there is no gold standard for the measurement of SVMC in children with CP available.
- The psychometric properties of only a few SVMC measures have been evaluated.
- There is a lack of responsive assessment tools for the evaluation of therapy-induced changes in SVMC in neurological patients.
- Therefore, it is not known which measurement technique is best, regarding its psychometric qualities, for testing SVMC in children with CP.

Therefore, in accordance with the bullet points above, this PhD included the following research questions.

Study 1

The aim of the first study was to investigate the influence and relevance of SVMC of the lower extremity in children with CP on the ICF activity level. We used regression modelling to determine the strongest predictor for gait capacity among the following UMN signs: impaired SVMC, spasticity, muscle weakness and impaired trunk control. Thereby, we formulated the following research questions:

Research Question 1: Is SVMC, within a model of other common UMN motor impairments (spasticity, muscle weakness, impaired trunk control), an important predictor for gait capacity in children with CP?

Research Question 2: What is the association of SVMC with other lower extremity impairments (spasticity, muscle weakness) and impaired trunk control?

Study 2

Secondly, in relation to the recently noticed importance of SVMC in children with CP, it was of great clinical and scientific interest to translate the SCALE into German and to test the psychometric properties of this translated version.

Research Question 3: Can the original version of the SCALE be translated into German, thus, by following international standardized translation procedure for such purposes?

Research Question 4: How strong is the construct validity of the SCALE's German version, in terms of 1) its correlation with the Fugl-Meyer test (criterion validity); 2) its correlation with the GMFCS (concurrent validity) and 3) its discriminative validity, by showing significant differences 3.1) across and between joint (pairs) (proximal-distal-concordance); 3.2) for limb involvement (less/more affected limb) and 3.3) across and between GMFCS levels.

Research Question 5: Is the inter-rater reliability of the German version of the SCALE as reliable as the original version?

Research Question 6: What is the intra-rater reliability of the German version of the SCALE?

Study 3

Parallel to study 2 a systematic review of the psychometric properties of SVMC measurement tools for the lower extremity in children with CP was carried out:

Research Question 7: Which SVMC measures exist for children with CP?

Research Question 8: What is the level of evidence for the psychometric properties of these measures?

Research Question 9: Which (if any) tool would be best suitable, regarding feasibility and psychometric quality, for measuring therapy-induced changes of SVMC in the lower extremity in children with CP?

Study 4

Based on the results of studies 2 and 3, the aim of the fourth study was to evaluate the psychometric properties of a new measurement tool for measuring SVMC of the

lower extremity in children with CP. This new outcome measure uses the SCALE testing procedures combined with sEMG recordings of 10 lower extremity muscles to calculate the so-called “Similarity Index”. It was developed to create a SVMC outcome measure which would be sufficiently valid, reliable and sensitive to measure therapy-induced changes in SVMC. Although the author is aware of the importance of the assessment of responsiveness of a new tool, this could not be realized within the scope of this thesis.

The following research questions determined the structure of this study:

Research Question 10: Can the criterion validity of the SCALE-SI be established by a high correlation with the original SCALE?

Research Question 11: Considering concurrent validity: Does the SCALE-SI show a similarly strong correlation with the GMFCS as the original SCALE?

Research Question 12: Considering discriminative validity: Do SCALE-SI scores differ significantly between 1) neurological intact children and children with CP; 2) joint levels of both the less and more affected limb; and 3) limb involvement (less/more affected SCALE-SI score)?

Research Question 13: Does the “SCALE-SI” show adequate relative and absolute test-retest reliability values?

Taking all these different study aims into consideration, the overall aim of this PhD study was to investigate how to measure SVMC in the lower extremity in children with CP. Only by using robust outcome measures, it will be possible to evaluate therapeutic techniques aiming to improve children’s SVMC. Consequently, this will enable therapists to give parents an evidence-based answer to questions like “Will my child learn to walk normally due to your therapy?” or “Can you teach him to walk properly/normally?”

Chapter 3: Study 1

3.1 Purpose of chapter

The purpose of this chapter is to describe the methodology and present and discuss the results in relation to the first two research questions (Section 2.4) concerning the relative contribution of SVMC and other UMN motor impairments as well as the variability in gait capacity observed in children with cerebral palsy.

This chapter has been published in the Journal Disability and Rehabilitation in 2017: Balzer J, Marsico P, Mitteregger E, van der Linden ML, Mercer TH, van Hedel HJ. 2017. Influence of trunk control and lower extremity impairments on gait capacity in children with cerebral palsy. *Disabil Rehabil.* 24:1-7 (Appendix 10).

3.2 Background

As described in chapter 2.2.7., although impaired SVMC is an impairment on the ICF body functions and structure level, it might cause limitations at ICF activity and participation level together with other common UMN signs. The importance of SVMC with regard to gross motor function has recently been shown in several studies (Chruscikowski et al. 2017; Vos et al. 2016; Park and Kim 2013; Kim and Park 2011; Ross and Engsberg 2002; Voorman et al. 2007; Desloovere et al. 2006) in ambulatory children with CP (GMFCS I-III). In these studies, SVMC and muscle strength were found to be stronger predictors for gross motor function compared to spasticity and range of motion (contractures). As the lower extremities and the trunk are involved during walking, knowledge about their individual and combined impact on gait is essential for developing optimal gait-rehabilitation intervention strategies (Chruscikowski et al. 2017, Saether et al. 2014). However, the influence of trunk control on gait received lesser attention – until now. Traditionally, disturbed motor control and lower extremity impairments were seen as the primary and secondary gait deviations, (Gage et al. 2009). Abnormality of trunk kinematics during walking was mostly considered as a compensatory gait deviation. Recently, this focus has been amended by an increasing number of studies investigating the trunk and upper limbs during walking (Saether et al. 2014; Heyrman et al. 2014; Attias et al. 2014). This shift in focus coincided with the development of new trunk control measures for children with CP (Heyrman et al. 2014; Attias et al. 2014). The first studies in this field provide increasing evidence that altered trunk control during gait in children with CP should

not be considered solely as a compensation for gait deviation (due to altered lower extremity functioning), but should also be considered as a direct aspect of gait deviation (Saether et al. 2014; Heyrman et al. 2014). Until now, studies of ambulatory function have investigated the contribution of either lower extremity impairments (Noble et al. 2017; Vos et al. 2016; Park and Kim 2013; Kim and Park 2011; Ross and Engsberg 2007; Voorman et al. 2007; Desloovere et al. 2006) or trunk control (Heyrman et al. 2014; Attias et al. 2014) with no study assessing both lower extremity impairments and trunk control, and evaluating their influence on gait. As both body functions are dependent on adequate motor control, the primary aim of this study was to investigate the impact of SVMC of the lower extremity and trunk on gait capacity (e.g. in this study, the time needed to perform the modified Timed Up and Go test or mTUG).

Although the importance of SVMC and trunk control on ambulatory function in children with CP has been investigated, the association between these two impairments and that with other lower limb impairments (i.e. spasticity, muscle weakness) has not yet been established. For instance, as described in chapter 2.2.6, it remains uncertain whether the pathophysiological nature of impaired SVMC is a negative UMN on its own (due to a loss of CST connections), just one functional feature of a positive UMN pathology (due to a loss of inhibition and the release of primitive flexor/extensor patterns), or a combination of both. As these different aetiologies may have different implications for treatment of gait disorders, gaining more insight in the interdependency of these UMN signs is of clinical relevance. Several studies have reported the association between SVMC with other UMN lesions: A fair correlation (ρ = Spearman's rho) between SVMC and spasticity of the plantarflexors (ρ = 0.300, p = 0.004; Østensjo et al. 2004) and with (ankle) range of motion (ρ = 0.303, p < 0.001; Park and Kim 2013); moderate correlation between SVMC and muscle strength (ρ = 0.505, p < 0.001; Park and Kim 2013) and muscle volume (ρ = -0.580, p < 0.001; Hanssen et al. 2018). However, studies investigating the association between motor control of the lower extremity and that of the trunk are lacking. In order to provide evidence-based clinical decision making while treating ambulatory children with UMN signs, the secondary aim of this study was to gain more knowledge with regard to the association between impaired SVMC and muscle weakness, spasticity, and trunk control.

Based on previous studies (Noble et al. 2017; Ross and Engsberg 2007; Voorman et al. 2007; Desloovere et al. 2006), the following *a priori* hypotheses were formulated: i) there is a negative, moderate (correlation coefficient $>.5$) relationship between gait capacity and SVMC, trunk control and lower extremity muscle strength; ii) there are positive, weak (correlation coefficient $<.5$) relationships between spasticity and gait capacity; iii) SVMC, muscle strength and trunk control are the strongest predictors for gait capacity; iv) there are a positive, moderate correlations between SVMC and muscle strength with trunk control and v) there is a negative, weak relationship between spasticity and trunk control.

3.3 Method

3.3.1 Participants

In- and out-patients of the “Rehabilitation Centre Affoltern am Albis, University Children’s Hospital Zurich” were recruited by convenient sampling. Inclusion criteria were: diagnosis of spastic CP, age between five and 20 years, ability to walk (Gross Motor Function Classification (GMFCS) level I-IV), and ability to follow simple instructions. Participants with additional movement disorders, with an unstable situation regarding their tonus-regulating medications and/or participants, who had a botulinum toxin injection within the last six months or any surgical correction within the last year, were excluded. The study was approved by the ethical committee of the Canton of Zurich (KEK-ZH-Nr.2011-0404). Informed consent and assent were obtained from parents and participants (respectively).

3.3.2 Measurements

All tests were carried out by the same two experienced neuro-pediatric physiotherapists within a maximum timeframe of one hour and in accordance to standardized procedures.

3.3.3 Lower extremity assessments

To assess SVMC at the hip, knee, ankle, subtalar and toe joints, the “Selective Control Assessment of the Lower Extremity” (SCALE), as described in section 2.3.6.1 of this thesis, was performed [27,28].

Spasticity and muscle weakness of hip, knee and ankle flexion and extension movements were assessed with the “Modified Ashworth Scale (MAS),” and the “Manual Muscle Test” (MMT), respectively.

The MAS (Bohannon and Smith 1987) scores spasticity on an ordinal scale ranging from “0” to “4” in accordance to the velocity dependent definition of spasticity from Katz et al. (1992). Although its criterion validity was established by using the pendulum test (Naghdi et al. 2007), its correlation with an increased alpha-motor-neuron activation (Fleuren et al. 2010), as well as with increased muscle activation and resistance (Fosang et al. 2003) ranged from weak to moderate only. In children with CP, interrater-reliability of the MAS for the lower extremity joints ranged from weak to good (Fosang et al. 2003; Mutlu et al. 2008).

Muscle strength was evaluated with the MMT in accordance to Kendall et al. (1993). Scores ranged from 0 – 5. Its scoring system was originally developed and tested on validity for determining muscle weakness in patients with poliomyelitis (Lovett and Martin 1916). Its interrater-reliability has not yet been tested in children with CP, but was moderately too good for children with muscular dystrophy (Cuthbert and Goodheart 2007). Although evaluation of the psychometric properties for the MAS and MMT in children with CP is limited, they were performed, as they are considered the clinical standard. Moreover, they have been used in previous studies, hence allow the comparison of our results (Vos et al. 2016; Ross and Engsberg 2007; Østensjø et al. 2004).

3.3.4 Trunk control assessment

Trunk control was assessed using the “Trunk Control Measurement Scale” (TCMS). This is a 15-item assessment that examines sitting balance during functional activities (Heyrman et al. 2011). The TCMS considers that the trunk should provide a stable base of support and that it is also an actively moving body segment. The first five items test static sitting balance, followed by ten items that test dynamic sitting balance. Dynamic sitting balance is further divided into two subscales, seven items testing ‘selective movement control’ and three items testing ‘dynamic reaching’. Its validity was supported for children with spastic CP by i) moderate to strong correlations with the “Gross Motor Function Measure” (GMFM) (Heyrman et al. 2011; Mitteregger et al. 2015), ii) significant differences between healthy children and children with CP (Heyrman et al. 2011) and iii) a strong correlation with centre of pressure measures

whilst sitting (Mitteregger et al. 2015). Its interrater-reliability was established as the ICC was 0.91 (Heyrman et al. 2011).

3.3.5 Gait capacity assessment

For assessing the participants' gait capacity, the paediatric version of the mTUG (Williams et al. 2005) was performed. It records the time a child needs to stand up from a chair with foot contact, to walk three meters to a target, turn around and return to the chair and finally sit down. Reliability and validity of the mTUG was supported by a study in a sample of 176 children without physical disabilities and 41 young people with physical disabilities due to CP or spina bifida (Williams et al. 2005). In this study, the mTUG was performed twice and the average time of the two trials was included into the analysis.

3.3.6 Statistical Analysis

Statistical analysis was performed with SPSS 17.0 (IBM, Armonk, USA). Alpha was set at 0.05 (two-tailed). The Shapiro-Wilk-test showed that the data of most scores was not normally distributed. Hence, Spearman's correlation coefficients ($\rho = 1 - (6\sum d^2) / n(n^2 - 1)$); Abbreviations: ρ = Spearman's Rank Correlation Coefficient; d = differences between the ranks assigned to each squared observation; n = the numbers of observations in each (both) data sets)) were calculated between the mTUG, age, SCALE, MMT, MAS and TCMS total and sub-scores. We also calculated ρ between the TCMS total and sub-scores and the SCALE, MMT and MAS scores.

In a second step, simple and multiple linear regression analysis (backward modelling) (Miles and Shevlin, 2001) were carried out to determine the most important predictor(s) for explaining mTUG variance. A model using SCALE, MMT, MAS and TCMS total scores as independent variables was analysed. For the regression analysis, the following assumptions were checked: i) homogeneity of variance via a nonsignificant Levin's test; ii) lack of multicollinearity by calculating the tolerance and variance inflation factor for each independent variable; iii) lack of autocorrelation by calculating the Durbin-Watson test, and iii) a lack of outliers (case-wise diagnostic) based on the values of Cook's and Mahalanobis' distance (Bowerman and O'Connell 2000; Field 2005). Based on the regression sample size guidelines by Miles and Shevlin (Miles and Shevlin 2001), we aimed for a sample size of 50 participants which

would be sufficient for a regression model with four predictors with a moderate to large effect size.

3.4 Results

Sixty-eight children with spastic CP gave informed consent for participation. Due to a lack of compliance (lack of motivation, concentration problems) or due to organizational issues (unavailable walking aids), 14 data sets were incomplete. As case-wise diagnostic for the regression analysis revealed that mTUG scores of two participants (GMFCS level IV) laid three standard deviations above the mean, these participants were classed as outliers and omitted from the analyses. Therefore, demographic and performance characteristics of 52 participants are presented in table 4. The 23 females and 29 males were on average 11 years and 9 months (SD 4 years 6 months) old. Twenty-two children had a GMFCS level I, 12 had level II, 16 level III and two level IV. Further clinical characteristics are presented in table 4.

Table 4: Study 1 - Participants' clinical and functional characteristics

	spastic CP n=52		
Measures	median	(IQR)	range
MMT total score (0-60)	44.0	(20)	20 - 60
SCALE total score (0-20)	12.0	(5.2)	0 - 19
MAS total score (0-48)	2.5	(4)	0 - 20
TCMS TCMS – static (0-20)	19.0	(5.2)	4 - 20
TCMS – selective (0-28)	14.0	(8.2)	1 - 26
TCMS - dynamic (0-10)	8.0	(4.2)	1 - 10
TMS - total (0-58)	41	(14.2)	6 - 56
mTUG (s.)	7.9	(5.5)	4.3 - 47.5
Age (yy.mm)	11.7	(7.6)	5.9 - 19.11

Abbreviations: MMT: Manual Muscle Test; SCALE: Selective Control Assessment of the Lower Extremity; MAS: modified Ashworth Scale; mTUG: modified Time Up and Go test; TCMS: Trunk Control Measurement Scale, SD: Standard Deviation; IQR: InterQuartile Range

3.4.1 Correlation analysis

Correlation results for the lower extremity impairments, TCMS and mTUG are summarized in Table 5. The MMT total scores showed the strongest relationship with both the mTUG and TCMS total score, closely followed by the correlations between

the total SCALE scores and the mTUG and TCMS total score. Lowest correlations were found between the MAS total scores and the mTUG or TCMS total score and its sub-scores. Only the correlation between age and gait capacity was weak and non-significant. Corresponding scatter plots are shown in Figure 7. Furthermore, MMT and SCALE correlated strongly (Table 5).

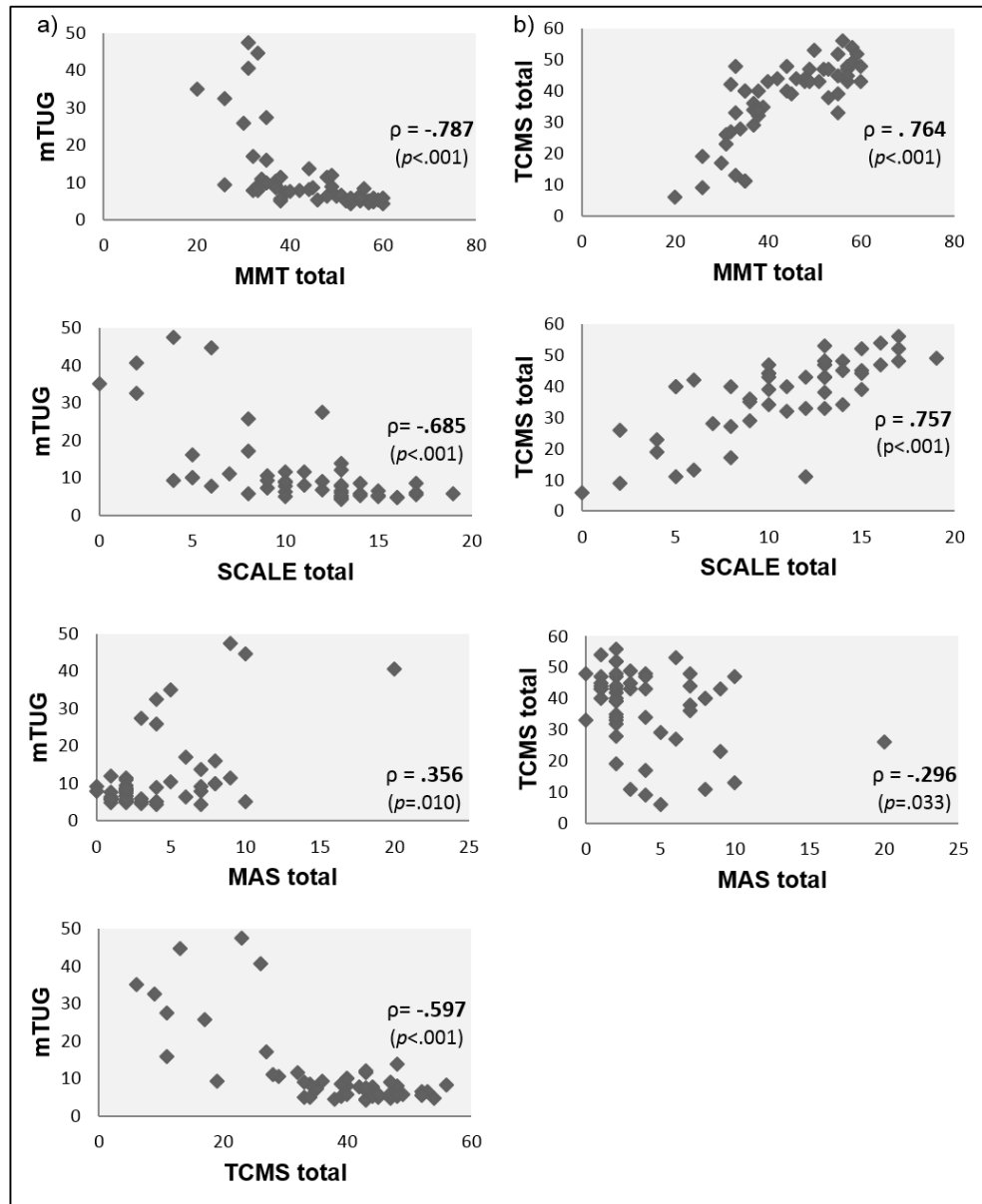


Figure 7: Study 1 - Scatterplots:

a) Scatter plots and Spearman's correlation coefficients (ρ) between lower extremity impairments, trunk control and gait capacity; b) Scatter plots and Spearman's correlation coefficients (ρ) between lower extremity impairments and trunk control

Abbreviations: MMT: Manual Muscle Test; SCALE: Selective Control Assessment of the Lower Extremity; MAS: modified Ashworth Scale; mTUG: modified Time Up and Go test; TCMS: Trunk Control Measurement Scale

Table 5: Study 1 - Spearman's correlation coefficients (ρ) between lower extremity impairments, trunk control and gait capacity

spearman`s rank (ρ)	MMT	SCALE	MAS	mTUG
MMT (total)	1.00	.849 ($p<.001$)	-.255 ($p=.068$)	-.787 ($p<.001$)
SCALE (total)	.849 ($p<.001$)	1.00	-.435 ($p=.002$)	-.685 ($p<.001$)
MAS (total)	-.255 ($p=.068$)	-.435 ($p=.002$)	1.00	.356 ($p=.010$)
TCMS - static	.711 ($p<.001$)	.604 ($p<.001$)	-.189 ($p<.001$)	-.695 ($p<.001$)
TCMS - selective	.665 ($p<.001$)	.717 ($p<.001$)	-.362 ($p<.001$)	-.493 ($p<.001$)
TCMS - dynamic	.770 ($p<.001$)	.675 ($p<.001$)	-.218 ($p<.001$)	-.614 ($p<.001$)
TCMS (total)	.764 ($p<.001$)	.757 ($p<.001$)	-.296 ($p=.033$)	-.597 ($p<.001$)
Age				.093 ($p=.126$)

Abbreviations: MMT: Manual Muscle Test; SCALE: Selective Control Assessment of the Lower Extremity; MAS: modified Ashworth Scale; mTUG: modified Time Up and Go test; TCMS: Trunk Control Measurement Scale

3.4.2 Simple and multiple linear regression analysis

The above shown strong correlation coefficients (≥ 0.75) between the SCALE, MMT and TCMS indicated a probable risk of multicollinearity between these three predictors. This threat of multicollinearity was further shown in the Variance Inflation Factor (VIF) values for the SCALE and MMT (VIF: 4.968 and 4.439, respectively). Also, the average VIF (3.38) of the model was greater than 1 and should be considered when interpreting the results of this regression analysis (Hair et al. 2017).

When applying a simple linear regression modelling to predict gait capacity, the TCMS total score alone explained most of the variance (54%) of the mTUG, followed by the SCALE (43%), the MMT (40%) and the MAS (31%). As age was not correlated with the mTUG, it was not included in the regression analysis (Table 6).

A multiple backward regression model was applied to investigate which lower extremity and/or trunk impairments explain the greatest amount of variance in gait

capacity. In the first step the total SCALE score was removed from the model, followed by the MMT score. The TCMS was the strongest predictor with a standardized regression coefficient " β " of -0.624 ($p < 0.001$), when explaining the variance in mTUG ($R^2 = .67$, $F_{(4,51)} = 49.246$, $p < 0.001$). Together with the MAS, the TCMS remained in the final model and they both explained overall 67% of the mTUG variance. To improve the interpretation of these findings, this analysis showed that a decrease of trunk control in the amount of 12 TCMS points resulted in a 6.6 seconds increase of the mTUG (Table 6).

Table 6: Study 1 - Simple and multiple linear regression analysis for predicting gait capacity

Dependent variable	Independent variable	B	Std. Error B	β	R^2
mTUG	simple linear regression				
	Constant	36.08	4.34		
	MMT (total score)	-0.56	0.09	-.637 ($p < .001$)	.40
	Constant	28.99	3.02		
	SCALE (total score)	-1.60	0.26	-.657 ($p < .001$)	.43
	Constant	4.93	1.87		
	MAS (total score)	1.67	0.35	.559 ($p < .001$)	.31
	Constant	34.93	3.21		
	TCMS total	-0.62	0.08	-.734 ($p < .001$)	.54
	Constant	11.83	4.32		
	Age	0.04	0.36	-.016 ($p = .909$)	.00
	multiple linear regression: MMT, SCALE, MAS, TCMS (backward modelling)				
	Step 1				
	Constant	26.91	4.34		
	SCALE	0.34	0.45	.139 ($p = .459$)	
	MMT	-0.20	0.16	-.226 ($p = .213$)	
	MAS	1.23	0.30	.410 ($p < .001$)	
	TCMS	-0.46	0.11	-.542 ($p < .001$)	.68
	Step 2				
	Constant	28.88	3.81		
	MMT	-0.12	0.12	-.136 ($p = .306$)	
	MAS	1.11	0.26	.371 ($p < .001$)	
	TCMS	-0.44	0.11	-.520 ($p < .001$)	.68 ($p = .459$)
	Step 3				
	Constant	26.92	3.31		
	MAS	1.12	0.26	.376 ($p < .001$)	
	TCMS	-.528	0.07	-.624 ($p < .001$)	.67 ($p = .306$)

Abbreviations: MMT: Manual Muscle Test; SCALE: Selective Control Assessment of the Lower Extremity; MAS: modified Ashworth Scale; mTUG: modified Time Up and Go test; TCMS: Trunk Control Measurement Scale; B: Beta (unstandardized regression coefficient); Std. Error B: standardized error of Beta; β : standardized regression coefficient; R^2 : coefficient of determination

3.5 Discussion

The current study showed that among measures of lower extremity and trunk impairment, trunk control was the strongest predictor for gait capacity in children with CP. Furthermore, the strong correlations between SVMC and muscle strength with trunk control as well as with gait, underpin the important influence of these impairments on walking capacity.

3.5.1 Prediction of gait capacity

Until now, no study has investigated the impact of both lower extremity impairments and trunk control on gait capacity. In deviation to the initial hypothesis, that trunk control and leg muscle strength would be the strongest predictors for gait capacity, the MMT (surprisingly) and the SCALE scores were excluded from the regression model. The unanticipated exclusion of the MMT, as well as the exclusion of the SCALE, is likely to be caused by multicollinearity between TCMS, MMT and SCALE. As multicollinearity is a methodological limitation of this study, its cause and consequences will be explained in further detail in the section 3.6 under limitations.

The results of the simple regression analysis are in agreement with those of previous studies, which reported the importance of SVMC (van der Linden et al. 2018; Chruscikowski et al. 2017; Park and Kim 2013; Kim and Park 2011; Ross and Engsberg 2007), strength (Park and Kim 2013; Kim and Park 2011; Ross and Engsberg 2007) and a minor influence of spasticity and gait capacity/performance (Park and Kim 2013; Kim and Park 2011; Ross and Engsberg 2007).

Nevertheless, a direct comparison regarding the absolute strength of the relationship between our study and previously published research, is not appropriate due to the existence of several methodological differences: i) previous studies used different dependent variables such as three dimensional gait analysis (Desloovere et al. 2006), gross-motor function (Vos et al. 2016; Voorman et al. 2007; Østensjø et al. 2004), ii) differences in assessments/methods were used to quantify lower extremity impairments, iii) different levels of GMFCS of the study population, and iv) different statistical analyses (i.e Park and Kim 2013).

Further comparing the simple and multiple regression results in terms of the importance of trunk control on gait, only one other study was found, which showed that trunk control (quantified by the “Segmental Assessment of Trunk Control”)

explained 38-40% variance of the GMFM in 92 children with CP (GMFCS I-V) (Curtis et al. 2014). In this study, the TCMS explained half of the variance of the mTUG within an ambulant sample (GMFCS I-IV). Although the results of this study confirmed the strong relationship between trunk control and gait capacity, a meaningful clinical interpretation of this finding in terms of causal relation between the two is difficult. This is due to the current lack of knowledge concerning the responsiveness of the TCMS and the lack of intervention studies which might have included the TCMS. Thereby, it is unknown how likely it is to increase a patient's TCMS score and whether this results in an improvement of the mTUG.

3.5.2 Relationship between lower extremity motor functioning and trunk control

Regarding the secondary objective of this study, the a priori formulated hypotheses were confirmed by Spearman rank correlation coefficients exceeding 0.7 for the MMT and SCALE with TCMS. However, the correlation between MAS and the TCMS was lower than expected. These outcomes seem to support two clinical impressions, formed prior to conducting this study. Namely that patients with better active trunk control (e.g. due to training) or passive trunk control (e.g. supported sitting or brace) have a better capacity for improving selective movements and strengthening of their lower extremities. Furthermore, the strong relationships between the trunk and the lower extremity functioning might be explained by their close neuroanatomical positions on Penfield's homunculus. To the author's knowledge, this was the first study, which investigated the relationship between these two motor-control dependent body functions. Only one recent study was found, which addressed a similar topic. Heyrman et al (2014) investigated the impact of lower leg kinematics on trunk deviations in children with CP assessed during walking (as opposed to when sitting, as in our study). For measuring lower limb movements, they used the Gait Profile Score (GPS). They found no significant correlations between the trunk parameters during gait (e.g. Trunk Profile Score) and the GPS ($r = 0.35$, $p = 0.13$) and only fair correlations between the TCMS and GPS ($r = -0.49$). Furthermore, the correlations between trunk parameters assessed during sitting (TCMS) and gait were higher ($r = -0.63$ - -0.43). Therefore, they suggested that trunk deviations during walking are not exclusively associated with the presence of lower limb gait impairments and can thus be regarded as a discrete source of impairment and not merely a compensation (Heyrman et al. 2014)

3.5.3 Limitations and Methodological Considerations

As mentioned above, the problem of multicollinearity should be considered when interpreting these results (Bowerman and O'Connell 2000). The correlation matrix revealed high correlations ($p > 0.6$) between the MMT vs. SCALE, MMT vs. TCMS, and SCALE vs. TCMS. Furthermore, the average variance inflation factor of the starting model was above "1", which is considered as a threat to the validity of the model (Field 2005). The presence of multicollinearity of the aforementioned variables makes it impossible to obtain unique estimates of the explained variance, because these variables account for the similar variance and their beta values are therefore interchangeable (Type II error) (Miles and Shevlin 2001).

The regression results showed that when predicting gait capacity by SCALE, MMT, MAS and TCMS scores, the SCALE and MMT scores were removed from the model since their scores explained a similar amount of variance in mTUG variance as the TCMS. Please note that these results do not indicate that SVMC and leg muscle strength do not influence gait capacity. The MAS, which on its own only correlates weakly with mTUG, seems to explain another part of the variance. Therefore, only the MAS and the variable with the highest beta value (TCMS) were kept in the final model.

This interpretation is supported by the results of the simple regression analyses, which showed that SCALE and MMT explain the second and the third largest amount of variance in mTUG (43% and 40%, respectively).

An alternative, although the more complex approach to handle multicollinearity, is to run a factor analysis (e.g. Structural Equational Modelling (Park and Kim 2013)) on the highly correlated predictors and to use the resulting factor scores (or latent variable) as a predictor (Hutchenson and Sofroniou 1999). As this statistical approach requires a larger sample size, it might be considered for future studies investigating similar research questions within a larger sample.

Concerning further methodological limitations about the generalizability of these study results, the dominance of participants with a higher gross motor/walking abilities level (GMFCS I and II = 66.4% versus GMFCS III and IV = 33.6%) should be considered. This underrepresentation of children with more severe mobility problems might also possibly explain the lower correlation with the MAS. The scatterplots of the MAS versus TCMS and mTUG reveal the dominance of participants with only a low level of spasticity.

Interpretation of the results is furthermore limited by the selected measures and joints/muscles. For instance, the focus of this study lied on flexors and extensors of the three major lower extremity joints. Furthermore, the participants' lower limb range of motion, which can potentially also affect a patient's gait capacity, was not reordered in this study. However, previous studies have shown no or weak correlations of this impairment with gait or walking performance (Desloovere et al. 2006; Østensjø et al. 2004). Moreover, when interpreting the results in relation to muscle strength, it should be considered that an isometric strength test (MMT) was used in order to allow for comparison with previous papers (Vos et al. 2016; Ross and Engsberg 2002; Voorman et al. 2007; Desloovere et al. 2006; Østensjø et al. 2004). An isokinetic strength measure might have been more appropriated in relation to gait capacity.

Although it would have been of clinical interest to know if particular joints or other assessments explain more variance than other joints or assessments, this would have required a much larger sample size. Therefore, based on the sample size guidelines (Miles and Shevlin 2001), only the total scores of the SCALE, MMT, MAS and TCMS were entered as predictors.

Another limitation of this study and those of previous studies (Vos et al. 2016; Ross and Engsberg 2002; Voorman et al. 2007; Desloovere et al. 2006; Østensjø et al. 2004) was that the chosen outcome measures (SCALE, MMT, MAS and TCMS) are rated by the observer rather than 'objectively' measured. To minimize a potential bias, the trained and experienced assessors (which were involved in previous studies on the translation, validation and reliability testing of these measures, see (Balzer et al. 2016; E. Mitteregger et al. 2015; P. Marsico et al. 2017; Petra Marsico et al. 2017)) used standardized protocols and, except for the first author, were unaware of the hypotheses of this study.

Finally, it should be considered that the mTUG (Williams et al. 2005) in comparison to other walking tests (i.e. 10-meter walking test) is a measure of both gait and balance activities and as such measures gait capacity in an ADL context. As the mTUG includes tasks like getting up, turning and sitting down, more leg strength and balance is required compared to walking only, children whose strength, balance and upper extremity functionality are more impaired (GMFCS levels III-IV) therefore experience considerably more difficulties with an mTUG than with straight walking only. This could have caused the non-linear relation between the SCALE and other measures of impaired body function and structures. A walking test, which only

assesses walking speed in a straight line (such as the 10-meter walking test) and not the ability to turn and to perform sit-to-stand transitions, may have resulted in more linear associations.

However, when searching for previous studies investigating the association between walking speed and measures of impaired body function and structures to underpin this assumption, only one study was found, which displayed scatterplots indicating the linearity of this association rather than correlation coefficients only. In this study (van der Linden et al. 2018), associations between lower limb impairment and RaceRunning speed were indeed more linear than the relationship observed in the current study.

3.5.4 Clinical implications and future work

Firstly, these study results show how SVMC and other common UMN impairments of the lower extremity and the trunk influence gait capacity, independently and/or in combination. In particular, the relevance of trunk control on gait capacity in children with CP was shown for the first time. Based on this finding and findings of previous studies investigating trunk control (Heyrman et al. 2014; Attias et al. 2014; Heyrman et al. 2011; Curtis et al. 2014), it is suggested that therapists should address the potential importance of trunk control in addition to lower extremity functioning when attempting to improve gait in children with GMFCS I-III.

Secondly, this study added new knowledge to the investigation on the interrelationships between of lower extremity and trunk motor control impairments. Results showed that while SVMC is strongly associated with muscle strength and trunk control, its association with spasticity is, although statistically significant, considerably weaker. This finding might advocate that impaired SVMC is a negative UMN sign by itself and not only a consequence of spasticity alone.

Thirdly, in line with previous studies, spasticity was not the most important limiting factor for gait capacity in children with CP (Vos et al. 2016; Ross and Engsberg 2002; Voorman et al. 2007; Desloovere et al. 2006; Østensjø et al. 2004). These findings as well as a recently increasing number of studies investigating the influence of SVMC on gait development (Chruscikowski et al. 2017) and gross motor function (Vos et al. 2016; Park and Kim 2013; Kim and Park 2011), challenge the traditionally claimed importance of spasticity management in ambulatory children with CP (GMFCS I-III).

3.6 Conclusion

This study aimed to increase the understanding of the two main issues. Firstly, the influence of impaired SVMC and other motor control impairments on gait capacity in children with CP were investigated. Secondly, the interdependency of these different motor impairments. By using a regression model to predict gait capacity, with lower extremity and trunk function as independent variables, it was shown that trunk control, muscle strength and SVMC account for a similar amount of variance in gait capacity variance. Spasticity accounted for the remaining, but considerably lower amount of mTUG variance. Correlation analysis revealed that impaired SVMC is less associated with spasticity in the current sample, advocating that it is a negative UMN on its own. Its strong correlation with gait capacity demonstrates its relevance for functional movements within this group of children. Overall, the results of this study may inform clinicians in their assessment and treatment of ambulatory children with CP.

Chapter 4: Study 2

4.1 Purpose of chapter

This chapter describes the translation as well as the assessment of the validity and reliability of the German version of the SCALE. Thereby the purpose of this chapter is to answer research questions three to six of Section 2.4.

This chapter was published in the Journal Developmental Medicine and Child Neurology in 2016: Balzer J, Marsico P, Mitteregger E, van der Linden ML, Mercer TH, van Hedel HJ. 2016. Construct validity and reliability of the Selective Control Assessment of the Lower Extremity in children with cerebral palsy. *Dev Med Child Neurol.* 58(2):167-72. (Appendix 11).

4.2 Background

The previous regression study (Balzer et al. 2017) on the influence of SVMC and other motor impairments on gait capacity in children with CP, underlined the importance of SVMC in daily motor control in this patient group. Furthermore, as explained in section 2.2.7.2, impaired selective activation can initiate and worsen a vicious cycle of limited active movement, joint contractures, hampered motor function and diminished activity, thereby causing pain and appearance of secondary deformities in children with CP (Chruscikowski et al. 2017; Vos et al. 2016; Park and Kim 2013; Kim and Park 2011; Ross and Engsberg 2002; Voorman et al. 2007; Desloovere et al. 2006). Therefore, robust assessment of SVMC in children with CP is of great clinical and scientific interest. Although the clinical importance of physiological muscle activation is obvious, routinely assessment of SVMC is rare in the clinical environment (Dobson 2010). As described in section 2.3.6.1, this lack of clinical SVMC assessments might be explained by the following main complications in measuring SVMC. Firstly, testing SVMC is challenged by the coexistence of other motor signs (e.g. muscle weakness, spasticity). Secondly, SVMC needs to be measured indirectly (Cahill-Rowley and Rose 2014). Thirdly, in the German speaking areas, clinical application of assessments in another language is rare due to the language barrier. To enable regular assessment of SVMC also in German speaking areas and to thereby improve our knowledge in this central impairment, the aim of this study was to translate the English SCALE version, which was originally developed by Fowler and colleagues in 2009. The SCALE was selected for translation for the following reasons: i) it allows to

measure SVMC in accordance to the above stated definition from Sanger et al. (2003), ii) it assesses five main joints of the lower extremity; iii) testing position and procedure are well selected for children with bilateral involvement; iv) SCALE's validity and reliability testing have shown promising results (Fowler et al. 2009; Fowler et al. 2010).

To establish discriminative validity of the German SCALE's version, the following hypotheses were set: i) the German SCALE version would show a similar proximal-distal concordance than the original version (Fowler et al. 2009), ii) SCALE scores would differ significantly between the less and more affected limbs, and iii) across and between different GMFCS levels. For concurrent validity, a high positive correlation ($p > 0.70$) between the SCALE and the Fugl-Meyer Assessment (FMA) was expected. Furthermore, it was hypothesized that children with spastic CP and a high degree of muscle weakness and/or spasticity would have lower SCALE scores. For reliability, ICC values exceeding 0.8 (see also Fowler et al. 2010) and acceptable levels of absolute measurement error were expected.

4.3 Method

4.3.1 Participants

In- and out-patients of the "Rehabilitation Centre Affoltern am Albis, University Children's Hospital Zurich" were recruited by convenience sampling. A minimum sample size of 25-30 participants was required, to provide an accurate estimate of the random error. Inclusion criteria were: diagnosis of CP, age between five and 20 years, ability to walk (GMFCS level I-IV), and the ability to follow simple instructions. Participants with an unstable situation regarding their tonus-regulating medications and/or who had a botulinum toxin injection within the last six months or any surgical correction within the last year were excluded. The study was approved by the ethical committee of the Canton of Zurich (KEK-ZH-Nr.2011-0404). Informed consent and assent were obtained from parents and participants.

4.3.2 Measurements

In order to translate the original SCALE version into German, the following international guidelines from Beaton et al. (2007) were followed: i) translation into German by two independent native German speaking physiotherapists, ii) creation of consensus version, iii) back-translation into English by a translation company, and iv)

endorsement by the authors of the original version (Beaton et al. 2007). The final German SCALE version is provided in Appendix 4. For comparison the original English Version of the SCALE can be found in Appendix 5.

Testing procedures were standardized according to the assessment guidelines. All tests were carried out by the same two experienced neuro-paediatric physiotherapists, one assessing and one assisting. Tests were performed for both legs within a maximum timeframe of one hour.

SCALE administration required patients to perform specific isolated movement patterns at the hip, knee, ankle, subtalar and toe joint. SVMC of each joint movement was scored on a three-point ordinal scale. SVMC was scored as “normal” (2 points) if the patient could move the tested joint isolated (e.g. without moving other joints), within at least 50% of the possible range of motion and at a physiological cadence cued verbally by the therapist (e.g. “flex, extend, flex”). If any deviation in performance occurred (movement performed slower, below 50% of the range of movement, with co-/mirror-/synergistic-movements), selectivity was regarded as “impaired” (1 point). The score “unable” was given, if no joint movement could be made or mass-synergy-patterns occurred. SVMC was scored separately for each joint, for each limb and for both limbs together.

To analyse discriminative validity, patients were classified according to their limb involvement and GMFCS level (I-IV). The MMT (Kendall et al. 1993) leg-score was used to determine the more and less affected leg. If MMT scores were similar, further differentiation was based MAS (Bohannon and Smith 1987) scores.

To assess the SCALE’s concurrent validity, the FMA (Woodbury et al. 2007) was measured. The FMA is a valid assessment tool for testing SVMC in stroke and contains specific items for testing selectivity of the knee (items III a; IVa) and ankle joint (items III b; IV b). Like the SCALE, the FMA uses a 3-point ordinal scale to score (0 = cannot perform; 1 = performs partly; 2 = performs fully) selectivity of the joint movement.

Furthermore, when correctly applied, the MMT should also reflect the selective activation of a muscle (group). Therefore strength of the hip and knee flexors and extensors and of ankle dorsi- and plantar-flexors were assessed by the MMT (0 to 5; (Kendall et al. 1993).

Despite spasticity and SVMC being different constructs, spasticity can negatively influence SVMC, wherefore we were interested in correlating the SCALE with MAS scores (0-4; (Bohannon and Smith 1987)). We assessed the MAS also for hip, knee and ankle joints.

The SCALE assessment was video-taped for (intra- and inter-rater) reliability testing, in order to minimize participants' strain. The camera was positioned in front of the participant. This position allowed the observation of the tested joint movement and possible compensatory and mirror movements of the contralateral limb as well as other body parts. Although an additional video from the sagittal plane may have allowed for a more accurate evaluation of the range of motion of the ankle and knee joint, none of the raters experienced difficulties in evaluating whether the movement exceeded 50% of the passive range of motion (one criterion that differentiates between normal or impaired SVMC) or not. For reliability testing, the videotaped assessment was scored twice within a timeframe of six to eight weeks after the first scoring. Rater(s) was (were) blinded to the results of the first scoring (intra-rater) or results from the other rater (inter-rater).

4.3.3 Statistical Analysis

The Shapiro-Wilk-test showed that most scores were not distributed normally, hence, non-parametric statistical tests were used.

Therefore, a Friedman-test was performed to determine whether SCALE scores differed between all joint-pairs of each leg. Alpha was set at 0.05 (two-tailed). Post-hoc differences between adjacent joints (e.g. hip versus knee), as well as between sum scores of the more and less involved leg, were determined with the Wilcoxon signed rank test (to adjust for multiple comparisons, alpha was set at 0.01). Differences in total SCALE scores for children categorized via GMFCS level were evaluated with the Kruskal-Wallis test. A priori post-hoc significance levels were set for successive pair-wise testing between adjacent GMFCS levels (e.g. level I versus II, II versus III) with Mann-Whitney-U tests (post-hoc tests: alpha=0.025). Effect size (r) was calculated in accordance to Rosenthal (1994) ($r = |Z| / \sqrt{N}$ | Abbreviations: Z = z-score; N = number of participants (sample)).

To further evaluate the validity of the SCALE Spearman's rank correlation coefficients ($\rho = 1 - (6\sum d^2) / n(n^2 - 1)$) between SCALE scores on joint, limb and total levels and

FMA, MMT and MAS scores were calculated. Relative intra- and inter-rater-reliability was evaluated ICCs 2.1 (two-way mixed model; type absolute agreement: $ICC(2,1) = \frac{MS_R - MS_E}{MS_R + (k-1)MS_E + (k/n)(MS_C - MS_E)}$; *Abbreviations:* MS_R = mean square for rows; MS_E = mean square for error; MS_C = mean square for columns; n = number of participants (sample); k = number of raters/measurements; (Koo and Li 2016)). Absolute reliability was determined by the SEM ($SEM = SD \times \sqrt{1 - r_{xx}}$; *Abbreviations:* *SEM*: Standard Error of Measurement, *SD*: Standard Deviation; r_{xx} : reliability value ($ICC(2,1)$) and the minimal detectable change (MDC) ($MDC = SEM \times \sqrt{2} \times 1.96$; *Abbreviations:* *SEM*: Standard Error of Measurement) (Haley and Fragala-Pinkham 2006). Statistical analysis was performed with SPSS 17.0 (IBM, Armonk, USA).

4.4 Results

Forty-two children with spastic CP gave informed consent to participate in this study. One child did not complete the assessments due to a lack of compliance. As an allocation of the more and less affected leg was not possible in two data sets, these datasets were omitted from all analyses. Therefore, demographic and clinical characteristics of 39 children with spastic CP (unilateral $n=20$; bilateral $n=19$) were available. Eighteen children were female. Twenty-three children had a GMFCS level I, five had level II, eight level III and three level IV. Further characteristics are presented in Table 7.

Table 7: Study 2 - Participants' characteristics

Measures	Spastic CP (n=39) mean age: 12y 6mo [SD 3y 7mo]		
	mean (SD)	median (IQR)	range
SCALE			
less affected leg (x/10)	6.6 (2.8)	7.0 (4.0)	0 - 10
more affected leg (x/10)	4.5 (2.0)	5.0 (3.0)	0 - 9
total score (x/20)	11.2 (4.5)	13.0 (4.5)	0 - 19
FMA			
less affected leg (x/8)	5.7 (2.9)	8.0 (4.0)	0 - 8
more affected leg (x/8)	3.8 (2.5)	5.0 (3.5)	0 - 8
total score(x/16)	9.6 (5.1)	11.0 (8.0)	0 - 16
MMT			
less affected leg (x/30)	24.8 (6.1)	28.0 (11.0)	10 - 30
more affected leg(x/30)	20.9 (5.5)	23.0 (9.0)	9 - 30
total score(x/60)	45.6 (11.0)	49.0 (17.5)	20 - 59
MAS			
less affected leg (x/24)	1.9 (3.1)	1.0 (2.5)	0 - 14
more affected leg (x/24)	3.4 (3.2)	2.0 (3.0)	1 - 16
total score (x/48)	5.4 (6.0)	3.0 (5.0)	1 - 30

Abbreviations: MMT: Manual Muscle Test; SCALE: Selective Control Assessment of the Lower Extremity; MAS: modified Ashworth Scale; mTUG: modified Time Up and Go test; TCMS: Trunk Control Measurement Scale, SD: Standard Deviation; IQR: InterQuartile Range

The German version of the SCALE can be found in Appendix 4.

4.4.1 Discriminative validity

SCALE scores of contra-lateral joint-pairs (e.g. knee vs knee) of the less affected leg were significantly higher compared to those of the more affected leg, with the exception of the hip joint (Wilcoxon Signed Ranks Test: $Z = -1.414$, $p = 0.157$, $r = |Z / \sqrt{N}| = .226$) (Figure 8a). SCALE scores were generally lower for distal compared to proximal joints for both legs (Friedman's-test, Chi-square value of 89,988 ($p < 0.001$) for the less affected side, Chi-square value of 129,036 ($p < 0.001$) for the more affected side), except of the ankle versus toes for the less affected leg and bilaterally for the subtalar joint on toes (Figure 8a). SCALE limb scores were higher for the less (median (Mdn)=7; interquartile range (IQR)=0-10) compared to the more affected limb (Mdn=5; IQR=0-9; Wilcoxon Signed Ranks Test: $Z = -4.649$, $p < 0.001$, $r = .744$). When classifying participants in accordance to their diagnosis, statistically significant differences between the less and more affected limb were present for children with unilateral limb involvement (less affected: Mdn=9; IQR=7-10 versus more affected:

Mdn=4.5; IQR=3-6; Wilcoxon Signed Ranks Test: $Z = -3.733$, $p < 0.001$, $r = .835$) and bilateral involvement (less affected: Mdn=6; IQR=2.5-6 versus more affected: Mdn=5; IQR=3.5-7; Wilcoxon Signed Ranks Test: $Z = -2.877$, $p = 0.003$, $r = .660$) (Figure 8b).

Furthermore, SCALE scores differed significantly between GMFCS levels (Kruskal-Wallis-test: Chi-square value = 12.058, $p < 0.001$), and more specifically between GMFCS level I and II (Mann-Whitney-U-test: $U = 13.500$, $p = 0.007$, $r = .506$) (Figure 8c).

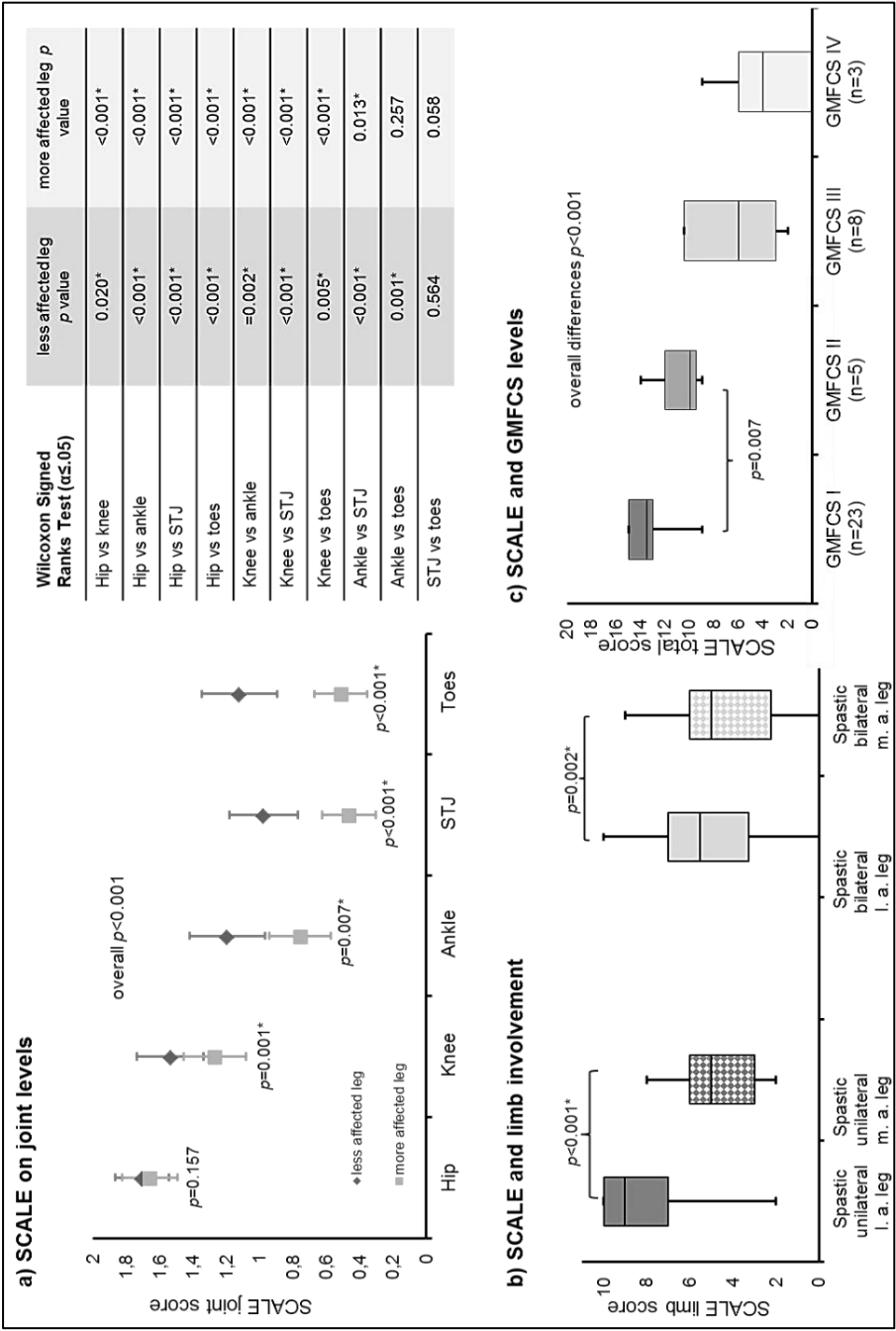


Figure 8: Study 2 - Discriminative validity for children with spastic CP:

a) SCALE and joint-pairs: SCALE joint scores between more and less involved limb and on adjacent pairs of joints: Friedman's-test and post-hoc Wilcoxon-test (post-hoc tests: p-values below 0.01 are considered significant) b) SCALE and limb distribution: total SCALE scores of less affected (l.a.) versus more affected (m.a.) leg in children with spastic unilateral and bilateral limb involvement: Wilcoxon-test c) SCALE and GMFCS levels: significant differences between all GMFCS levels (Kruskal-Wallis-test) and between GMFCS I versus II (post-hoc Mann-Whitney-U-test: p-values below 0.025 are considered significant)

Abbreviations: SCALE: Selective Control Assessment of the Lower Extremity; GMFCS: Gross Motor Function Classification System; l.a.: less affected; m.a.: more affected

4.4.2 Correlations

For the total SCALE score high correlations between and FMA and MMT were found (Figure 9a). The magnitude of the correlations between SCALE limb and joint scores and the clinical measures were comparable to those presented for the total scores.

There was a negative moderate correlation between the SCALE and the MAS total scores (Figure 9b).

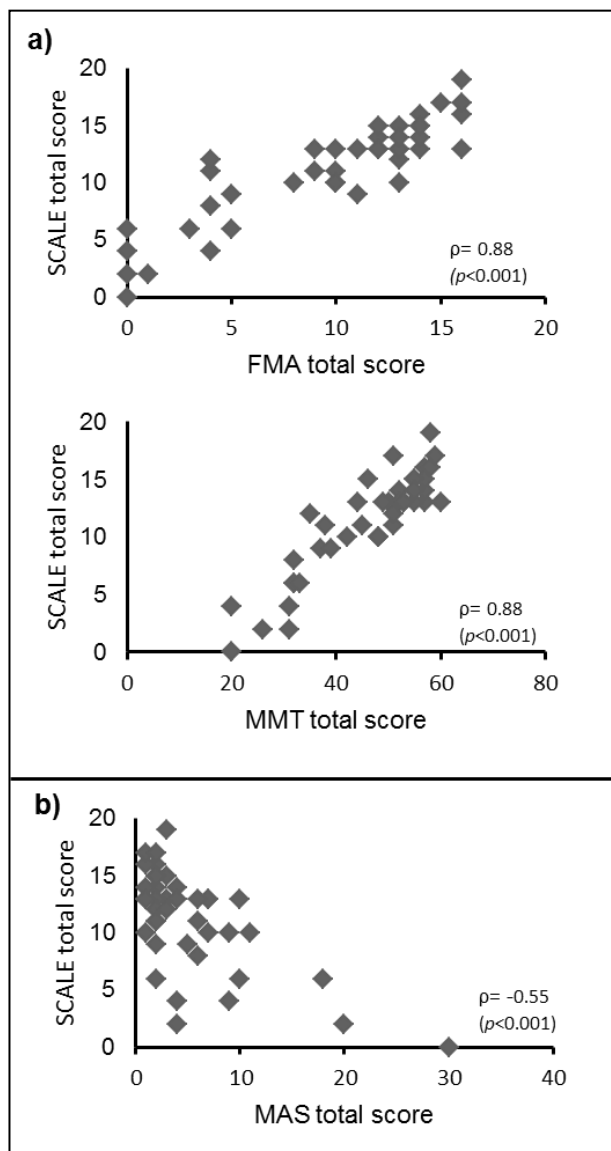


Figure 9: Study 2 – Spearman's rank correlation coefficient (ρ) of total SCALE scores and common clinical assessments for children with spastic CP:

a) concurrent validity: SCALE versus Fugl-Meyer Assessment (FMA) and SCALE versus Manual Muscle Test (MMT); b) correlation SCALE and Modified Ashworth Scale (MAS)

Abbreviations: SCALE: Selective Control Assessment of the Lower Extremity; FMA: Fugl-Meyer Assessment MMT: Manual Muscle Test; MAS: modified Ashworth Scale

4.4.3 Reliability

With ICC values exceeding 0.9 for limb and 0.8 for joint SCALE scores in children with spastic, intra- and inter-rater-reliability of the SCALE can be considered excellent. The MDC varied between 1.79 and 1.96 points (Table 8).

Table 8: Study 2 - Intra- and Inter-rater reliability of the SCALE

		Spastic CP (n=38)			
		Intra-rater reliability		Inter-rater reliability	
		less affected leg	more affected leg	less affected leg	more affected leg
Descriptive	Mean (SD)1	6.55 (2.86)	4.63 (2.16)	6.55 (2.86)	4.63 (2.16)
	Mean (SD)2	6.00 (2.81)	4.74 (2.50)	6.29 (2.94)	4.53 (2.16)
Relative Reliability	ICC	0.92	0.96	0.94	0.91
	p value	<0.001	<0.001	<0.001	<0.001
	95% CI	0.90 - 0.97	0.93 - 0.98	0.89 - 0.97	0.84 - 0.96
Absolute Reliability (SCALE points)	SEM	0.71	0.64	0.69	0.68
	MDC ₉₅	1.96	1.79	1.92	1.88

Abbreviations: ICC, Intra-class correlation coefficient; CI, Confidence Interval; SD, Standard Deviation; SEM, Standard Error of Measurement; MDC₉₅, Minimal Detectable Change at 95% confidence. Please note, due to a failure in a video-recording of the SCALE, we could include data from only 38 participants in the reliability analyses.

4.5 Discussion

Construct validity as well as intra- and inter-rater reliability of the German SCALE version in children with spastic CP are supported by this study.

4.5.1 Validity

Regarding SCALE's discriminative validity between adjacent or contra-lateral joint-pairs, between more versus less affected limb and between GMFCS-levels, the hypotheses of this study were partly confirmed and in line with previous results (Fowler et al. 2009).

Concerning the SCALE's ability to discriminate between the more and less affected limb in children with hemi- and diplegia, differences were significant for both groups, but for the later subgroup the difference was below the MDC.

On joint-level, SCALE scores of the hip joint did not differ between the more and less impaired limb. This could be due to the limited number of participants with greater motor impairment at the hip (e.g. GMFCS level IV) and the large number of children

with near maximal scores (ceiling effect: ceiling effect, e.g. 10 participants had a maximum total SCALE score for their less affected leg). SCALE scores at most distal joint-pairs tended to be lower with an exception for comparison between the subtalar joint and the toes. This trend was observed previously by Fowler et al. (2010) and Brunnstrom (1966) who reported that selective inversion and eversion were described as the most challenging movements for children with CP as well as adult stroke patients. As these movements rarely occur in isolation during daily activities (but frequently in combination with the movement of other foot-joints in supination or pronation), their movement performance might be experienced as unusual. Another neurophysiological explanation might be that the cortical representation of the lower extremity is largest for the great toe (Rothwell 1987).

The SCALE's discriminative validity was reflected in an overall difference between the GMFCS levels. Nevertheless, due to the small sample size, there were only significant differences between GMFCS level I and II and interpretation should be handled with caution. Performing a power analysis (80% power, two-tailed alpha .05) revealed that a sample size of 19 participants in GMFCS level II and III and 29 participants in GMFCS level III and IV, would be required to determine statistically significant differences between these GMFCS levels.

The strong correlation between the SCALE and FMA, illustrated in figure 9a, confirms that both assessments measure broadly similar constructs. Correlation coefficients between SCALE and MMT were also of similar magnitude. This could indicate that a correctly applied MMT will partially reflect the ability to selectively activate a muscle (group).

Regarding the additional hypothesis in relation to the association between SVMC and spasticity, only a moderate negative correlation was found. This result is in line with the findings of the previous regression study. Specific in this sample was the large variation of SCALE scores in participants with low MAS scores (Figure 9b). This range of SCALE scores in children with low spasticity might indicate that a mild level of spasticity does not necessarily affect SVMC negatively, while a clearer inverse relationship between SCALE and MAS is seen in participants with higher MAS values. However, the latter would have to be confirmed in studies including participants with a larger range of MAS values than reported in our study in which the majority was only mildly affected (e.g. mostly GMFCS levels I and II).

Furthermore, like Fowler et al. (2009), we found a high inverse relationship ($\rho \leq -0.80$) between the severity of CP (GMFCS levels) and the total SCALE score. Tying this together with the results of the previous regression study, which showed that the SCALE strongly correlated with gait capacity (mTUG), the importance of SVMC for the child's mobility level was furthermore shown.

4.5.2 Reliability

The hypothesis regarding reliability of the SCALE was confirmed. Excellent intra- and inter-rater reliability for the SCALE could be established as well as clinically acceptable values of absolute reliability for SCALE limb scores. As these results are based on a second rating of video-recordings, the ICC's might be slightly higher than when rated via a second assessment, where interfering factors like the participant's compliance or state of health might have altered testing conditions. For future studies, the accuracy of the video recordings could be improved by performing an additional video recording from the sagittal plane. However, the current values of this study are similar to previously reported observations by Fowler and colleagues (2009).

4.5.3 Limitations and Methodological Considerations

With regard to the methodological limitations of this study, it should be mentioned that grouping the results for the less and more involved limb might have decreased variability between participants, which could have resulted in lower correlation coefficients and ICC values. Nevertheless, our observations are broadly comparable with previous reports (Fowler et al. 2009).

4.5.4 Clinical implications and future work

By establishing validity and reliability properties of the German SCALE version, this study contributed to the translated version of the SCALE to the German speaking medical field. This enables regular clinical assessment of SVMC of the lower extremity in children with CP and might, in the long term, improve medical understanding and treatment of this central impairment.

In order to improve our pathophysiological understanding of the "real" nature of SVMC, especially in relation to its connection with spasticity, future cohort-studies including participants with also higher levels of spasticity and lower mobility levels (GMFCS III-V) would be informative. Furthermore, in order to gain clinical meaningful

SCALE threshold values for discriminating between GMFCS levels, future studies with a large sample might want to include a 'ROC curve analysis'.

In relation to the absolute reliability of the SCALE, it was shown that an increase of more than 2 SCALE points for the more affected leg in children with CP could be considered a true change (MDC). A future study on the responsiveness of the SCALE is needed to provide insight into whether such changes can be achieved with current rehabilitative (e.g. training or botulinum toxin) or surgical interventions (e.g. SDR). In order to expand SCALE's application from a diagnostic tool to an evaluative one, and to counteract the limitation of its ordinal scale, it might be appropriate to consider combining its test procedure with a neurophysiological measure (Dobson 2010; Zwaan et al. 2012).

4.6 Conclusion

In conclusion, previous results about the SCALE's validity, inter-rater reliability and increased distal impairment of SVMC were supported by this study. New evidence for construct validity of the German SCALE version in relation to common clinical tests in children with spastic CP, as well as important reliability aspects such as intra-rater reliability and MDC values, were added. Future studies need to investigate SCALE's responsiveness and thereby evaluate its sensitivity in measuring therapy-induced changes of SVMC of the lower extremity in children with CP.

Chapter 5: Study 3

5.1 Purpose of chapter

The purpose of this chapter is to describe the methodology and to present and discuss the results of the systematic review evaluating psychometric properties of SVMC measures. Thereby it aims to answer research questions 7-9, looking for a valid, reliable and responsive tool for measuring therapy-induced changes of SVMC in children with CP.

This chapter was published in the Journal Developmental Medicine and Child Neurology in 2017: Balzer J, van der Linden ML, Mercer TH, van Hedel HJ. 2017. Selective voluntary motor control measures of the lower extremity in children with upper motor neuron lesions: A systematic Review. *Dev Med Child Neurol.* 59(7):699-705. (Appendix 13).

5.2 Background

In accordance to the previous study (study 2) (Balzer et al. 2016), the following two main results can be summarized in relation to the German SCALE version. Firstly, validity and reliability of the SCALE as an assessment tool for children with CP was confirmed. Secondly, the SCALE's limitation in relation within its ordinal-scaling in measuring therapy-induced changes of SVMC were discussed. As the SCALE was originally developed for diagnosis and prognostic purpose only, its ordinal scoring system might be too broad to measure clinical changes of SVMC. That improvements of SVMC are desirable was indicated by the results of the first study (Chapter 1, Balzer et al. 2017), showing that the importance of SVMC on gait in children with CP. Therefore, a systematic review of the psychometric properties of SVMC measures for the lower extremity for children with UMN lesions was carried out. Although the population of interest for this PhD project is children with CP, for the purpose of the review it was expanded to children with all UMN lesions in order to find all available SVMC outcome measures within the neuro-paediatric population.

As this was, at least to the author's knowledge, the first review of SVMC outcome measures, they were two aims of this review. Firstly, it was aimed to provide an overview of all available SVMC assessment tools for the lower extremity. Secondly, the level of evidence of the psychometric properties of the identified assessment tools

were evaluated. Thereby, the review was aimed to find (if any such tool is available) the SVMC outcome measures, which scored best on psychometric properties and clinical utility in children with a UMN lesion. The ultimate aim was to identify a measure of SVMC with the highest level of evidence for its psychometric properties and the best clinical utility.

5.3 Method

5.3.1 Search Strategy

In order to identify studies assessing the psychometric properties of outcome measures of SVMC in children with an UMN lesion, the following databases were searched without any time limit until July 2016: MEDLINE, EMBASE, CINAHL, PsycINFO, SCOPUS, Cochrane and PEDro. The search strategy included keywords and synonyms for SVMC, as well as names of tools previously used to measure SVMC, the population of interest, and a validated search filter for finding studies on measurement properties (Terwee et al. 2009). Please see Appendix 6 for the applied search strategies for all searched databases. In addition, the reference lists were hand-searched for articles to identify additional studies.

5.3.2 Study Selection

A previously developed proprietary database was used (Microsoft Access 2010) to systematically enter the data and score the methodological quality of the studies (Ammann-Reiffer et al. 2014). Inclusion and exclusion criteria were defined in advance. In accordance to the definition of SVMC stated in the introduction, only papers dealing with selective movement of one joint of the lower extremity or with a primary selective (not synergistic) voluntary multi-joint movement were included. For example, papers dealing with the ankle dorsiflexion during initial contact or investigating pathological synergy patterns during walking (e.g. activation of the m. rectus femoris and m. semitendinosus during swing phase) were included, whereas papers measuring SVMC over the whole gait cycle or during gross motor coordination tasks were excluded. Considering that SVMC comprehends how accurately and smoothly someone can isolate the selection of a particular muscle group, papers describing the measurement of submaximal torque steadiness were included (e.g. ICF body function level b7300 power of isolated muscle activation). However, studies on maximal voluntary contraction were excluded, as patients with impaired SVMC

tend to produce maximal force by using mass synergy patterns. Furthermore, neuroimaging measures, testing structural and metabolic intactness of the involved underlying neurophysiological structures, or networks involved in SVMC (e.g. ICF body structure levels 1100 CST, primary cortex) were excluded. Only papers dealing with children and youths (3 to 21 years of age) with UMN lesions were included. This age range was chosen for neurophysiological reasons (e.g. maturation of the corticospinal tract) and practical reasons (e.g. compliance/understanding). Studies with the explicit aim to assess one or more psychometric properties were included, as well as cohort-studies indirectly investigating the psychometric characteristics of an outcome measure by for instance looking at the difference between neurological intact children and those with UMN lesion. All other forms of indirect evidence (e.g. intervention studies) were excluded. Only manuscripts published in English and German were included for review.

Two reviewers (J.B. and M. vdL.) independently screened all titles and abstracts of the papers. In cases of doubt, the full text article was consulted to decide whether or not the study met the inclusion criteria. If no consensus could be achieved, there was a third reviewer available.

5.3.3 Quality Evaluation

Evaluation of the methodological quality of the included papers was carried out independently by J.B. and M. vdL. by using the 4-point rating scale ('excellent', 'good', 'fair', 'poor' or 'not applicable') of the COSMIN checklist (Mokkink and Terwee 2010). As explained in Section 2.3.2, the COSMIN-checklist consists of three domains, namely validity, reliability and responsiveness (Mokkink and Terwee 2010). Each domain contains one or more measurement properties. The reviewer selects the measurement properties (COSMIN boxes) evaluated in the study and scores the specific item-lists via the aforementioned ordinal scoring system. The lowest score of all items of the chosen COSMIN box determines the overall methodological quality of the paper. In line with previous COSMIN reviews in the field of neuro-paediatrics, we adopted the overall COSMIN score by omitting the item regarding sample size (Ammann-Reiffer et al. 2014; Gerber et al. 2016).

To ensure that both raters scored the papers in accordance to the guidelines and to allow other raters to arrive at the same conclusion, the following procedures were established prior to the independent COSMIN rating: raters familiarized with the

COSMIN manual and terminology and discussed the scoring of two papers and established additional rating rules (Appendix 7). Although the COSMIN manual provides general rules for all boxes and items, for some items, the COSMIN rating is still open to subjective interpretation, e.g. 'time interval appropriate'. It is for this reason that COSMIN itself recommends specification of additional rules for individual reviews (De Vet et al. 2011). If the two reviewers could not agree on a scoring, a third reviewer was available.

For the assessment of the quality of the measurement properties, the updated criteria suggested by Terwee et al. (2007) were applied (Appendix 8). The overall level of evidence for each SVMC measure and each measurement property was evaluated according to the Cochrane Back Review Group Criteria 'strong', 'moderate', 'limited', 'conflicting', 'unknown' (Appendix 9) (Van Tulder et al. 2003). This overall score was given in relation to the methodological quality of the study and the results of the measurement properties. Again, criteria for sample size were adapted as follows: sample size > 100 subjects of the combined studies was rated as 'strong' [+++ or ---]; sample size between 50-99 was rated as 'moderate' [++ or --]; sample size between 25-59 as 'limited' [+ or -]; and sample size fewer than 25 as 'unknown' [?](Ammann-Reiffer et al. 2014).

5.4 Results

5.4.1 Description of the included studies

The systematic search resulted in 3590 references being identified. Based on the titles and abstracts, 33 papers were included for full-text reading. After applying the inclusion and exclusion criteria, 17 papers were retained for review (Figure 10).

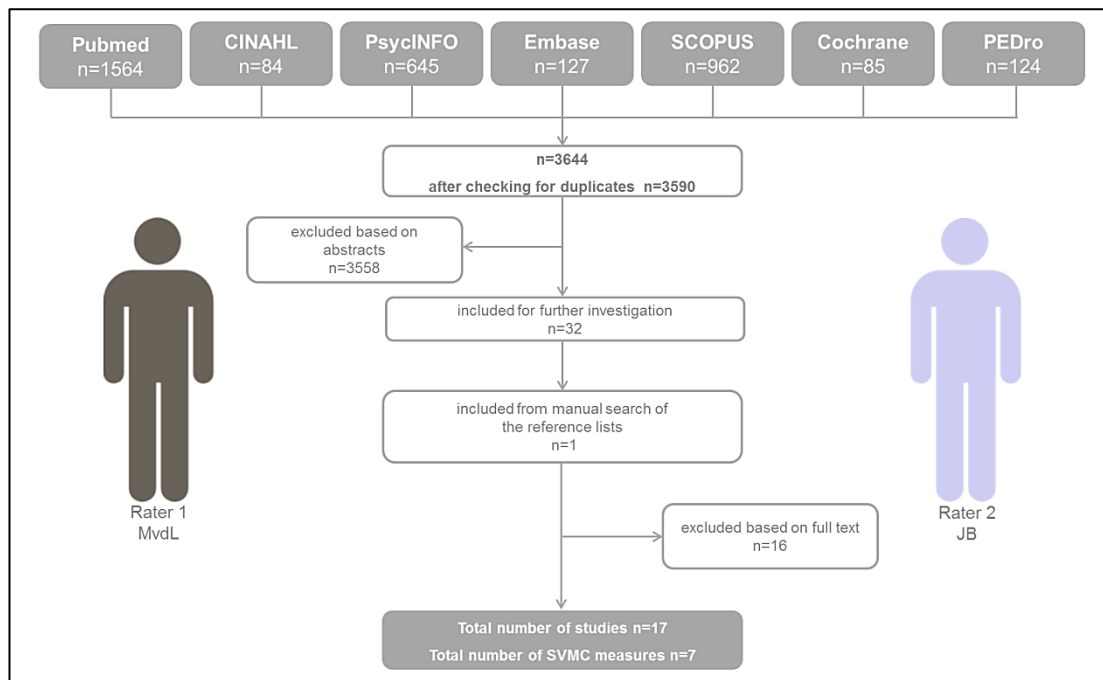


Figure 10: Study 3 - Flow chart systematic review

These 17 papers described the measurement properties of four clinical, ordinal-scaled, assessment tools (SMC; mTrost; Gillette's SMC test; SCALE) and three laboratory-based interval-scaled measurement tools (kinematic measures, sEMG, and torque steadiness). The majority of studies tested SVMC of the ankle or the knee joint.

The following psychometric properties were evaluated: hypotheses testing/construct validity was assessed in 17 studies, reliability in six (inter-rater $n=5$; test-retest $n=3$; intra-rater $n=1$) and both content and criterion validity in one study. Responsiveness was not evaluated in any study. Most studies tested the SCALE ($n=9$), followed by studies evaluating torque steadiness measures ($n=2$), kinematic measures ($n=4$), and sEMG of selected lower limb muscles ($n=1$). The age of the participants in the studies included for the final review ranged from two to 21 years, with the exception of one study in which the oldest participant was 28 years old. Although this age range was slightly wider than the one set by the inclusion criteria (3-21 years), discussing this issue ended in the common decision for inclusion. As the main age ranged from 9 years and 3 months to 16 years, the youngest and the oldest participant were seen as outliers. Sample size varied from eight to 51 participants. All studies included children with a diagnosis of cerebral palsy. In two cohort studies, data of children with CP who had undergone a SDR were compared to those of a control group children

with CP who had not (Engsberg et al. 2004). The comparison of SVMC between children with and without SDR was therefore considered an assessment of the validity of the SVMC tool. Four other cohort studies (Manikowska et al. 2016; Anon n.d.; Bandholm et al. 2009; Arpin et al. 2013) investigated the construct validity of the SVMC instrument, by comparing patients with CP versus participants who were neurologically intact. General characteristics and clinical utility for each SVMC measure is summarized in Table 9. The methodological quality per measurement property as well as the overall evidence criteria can be seen in Table 10.

5.4.2 Hypotheses testing

Of the 17 papers, which evaluated construct validity ('hypotheses testing'), ten papers included clinical assessment tools and seven papers laboratory-based measurement tools (Table 9 and 10)

Nine of the ten papers regarding clinical assessment tools evaluated construct validity of the SCALE. The modified COSMIN scores of three of these eight SCALE papers were 'good' (Fowler and Goldberg 2009; Fowler et al. 2010; Goldberg et al. 2011), four were rated as 'fair' (Fowler and Goldberg 2009; Rha et al. 2016; Rha et al. 2015; Lim 2015), one as 'excellent' (Balzer et al. 2016) and one as 'poor' (Yasuaki Kusumoto et al. 2016). Quality of construct validity was evaluated in accordance to Terwee et al. (2007) as 'positive [+]' in eight papers (Fowler et al. 2009; Fowler et al. 2010; Goldberg et al. 2011; Fowler and Goldberg 2009; Rha et al. 2015; Rha et al. 2016; Lim 2015; Balzer et al. 2016) and as mixed 'positive/negative [+]/[-]' in one study (Yasuaki Kusumoto et al. 2016). Overall, there was 'moderate positive [++]' evidence (Van Tulder et al. 2003) for construct validity of the SCALE in terms of: i) its correlation with the GMFCS (Fowler et al. 2009; Balzer et al. 2016) and ii) its proximal-distal concordance (SVMC is more often and/or more severely impaired in distal body parts) (Fowler et al. 2010; Balzer et al. 2016). A 'limited positive [+]' evidence level was given for its validity testing with the Berg Balance Scale (Lim 2015) and for predicting knee flexion during initial contact during stance phase of gait (Rha et al. 2016). Three studies investigated relationships between total limb SCALE scores and knee flexion during swing phase. Two studies (Goldberg et al. 2011; (Fowler and Goldberg 2009) found significant correlations, one did not (Rha et al. 2015). Therefore, their level of evidence was rated as 'conflicting [±]'. In relation to the poor quality of the Kusumoto et al. (2016) study, who investigated the relationship between the SCALE and knee

extensor strength, its level of evidence was rated as 'unknown [?]'. Therefore, this study did not contribute to the overall evidence level of the SCALE."

In the other three SVMC clinical assessment tools, construct validity was only evaluated for Gillette's SMC test. The study of Manikowska and colleagues (2016) compared Gillette's knee flexion SMC scores in patients with CP versus participants, who were neurologically intact using electromyography. This study received a 'poor' modified COSMIN score, but results were rated as 'positive [+]' in accordance to Terwee et al. (2007). The level of evidence (Van Tulder et al. 2003) was evaluated as 'unknown [?]'.

Seven papers investigated construct validity of SVMC using laboratory-based measurement tools including lower limb kinematics (Engsberg et al. 2004; Engsberg et al. 2008; Goldberg et al. 2011; Fowler and Goldberg 2009), sEMG (Zwaan et al. 2012) and torque steadiness (Bandholm et al. 2009; Arpin et al. 2013). Two kinematic papers (Goldberg et al. 2011; Fowler and Goldberg 2009) were already evaluated in relation to the SCALE's construct validity in the previous paragraph. As there is no gold standard measure for quantifying SVMC, and the papers are cohort studies investigating the correlation between the SCALE and kinematic measures for SVMC, they could be regarded as studies investigating the construct validity of the SCALE, but also of the kinematic measures, depending on which measure is regarded as more 'established'. As the quality and evidence rating of the two studies (Goldberg et al. 2011; Fowler and Goldberg 2009) is the same as presented above, the results will not be repeated here. Their COSMIN methodological quality was rated as 'fair' and quality of construct validity was rated as 'positive [+]' in three studies (Engsberg et al. 2008; Bandholm et al. 2009; Arpin et al. 2013) and as mixed 'positive/negative [+]/[-]' in two (Engsberg et al. 2004; Zwaan et al. 2012). As the sample size was too small ($n < 25$) in three studies, the evidence level was only scored for one kinematic (Engsberg et al. 2008) and one sEMG (Zwaan et al. 2012) study. As the results of these studies were ambiguous in supporting the construct validity of the SVMC measurement method, a 'conflicting [\pm]' evidence rating was assigned.

Overall, the methodological quality of the majority of the above mentioned studies was reduced due the absence of a priori formulated hypotheses, thereby limiting their COSMIN (Mokkink and Terwee 2010) as well as validity quality (Van Tulder et al. 2003) scoring.

5.4.3 Content and Criterion Validity

Content and criterion validity were only assessed for the SCALE (Table 10) (Fowler et al. 2009; Balzer et al. 2016) COSMIN rating (Mokkink and Terwee 2010) of content validity was considered 'poor'. Although 14 experts were involved on item-agreement for statements about content, administration and grading of the SCALE, the paper lacked a description whether all items were relevant for the construct or for the population of interest (Fowler et al. 2009). The quality rating of the results (Mokkink and Terwee 2010) was scored as 'indeterminate [?]', and the evidence as 'unknown [?]'.

The method applied to establish criterion validity of the SCALE was rated as 'excellent' (Balzer et al. 2016). As the Fugl-Meyer Assessment (item III and IV) measures a similar construct as the SCALE and their correlation exceeded 0.70, the SCALE criterion validity results were rated as 'positive [+]'.

Therefore, a 'limited positive [+]' evidence level was given for criterion validity of the SCALE.

5.4.4 Reliability

Reliability was investigated in three of the four clinical assessment tools (SMC (Löwing and Carlberg 2009; Smits et al. 2010), mTrost (Smits et al. 2010), SCALE (Fowler et al. 2009; Balzer et al. 2016)) and in two of five laboratory-based SVMC tools (kinematic (Engsberg et al. 2004), torque (Bandholm et al. 2009)). The SMC test-retest reliability was tested in two studies (Löwing and Carlberg 2009; Smits et al. 2010). The modified COSMIN rating for SMC inter-rater reliability ranged from 'fair' (Smits et al. 2010) to 'good' (Löwing and Carlberg 2009). The methodological quality of inter-rater reliability of the mTrost test was also rated 'fair' (Smits et al. 2010). Inter-rater reliability of the SCALE was tested by two studies and scored as 'excellent' (Balzer et al. 2016) and 'good' (Fowler et al. 2009). The SCALE's intra-rater reliability was further investigated and received a 'good' modified COSMIN score (Balzer et al. 2016). The methodological quality of test-retest reliability for the kinematic (Engsberg et al. 2004) and torque steadiness (Bandholm et al. 2009) measure was evaluated as 'good'. Overall, studies assessing inter-rater reliability were rated lowest for COSMIN items describing the statistical procedures (e.g. description of weighted scheme ICC, Kappa). The items regarding the stability of participants between the two or more assessments and the description of test conditions were the most limiting items for the four studies on test-retest and intra-rater-reliability.

Applying the quality criteria (Terwee et al. 2007) for measurement properties revealed 'positive [+]' results for four reliability studies (Engsberg et al. 2004; Fowler et al. 2009; Balzer et al. 2016; Löwing and Carlberg 2009); mixed 'positive/negative [+]/[-]' results in three studies (Bandholm et al. 2009; Löwing and Carlberg 2009; Smits et al. 2010), and negative '[-]' results for inter-rater reliability of the SMC test (Smits et al. 2010).

When evaluating the overall evidence level using Cochrane guidelines (Van Tulder et al. 2003), we found 'moderate positive results [++]' for the inter- and intra-rater reliability of the SCALE (Fowler et al. 2009; Balzer et al. 2016). 'Moderate negative results [- -]' were evident for the inter-rater reliability of the SMC (Löwing and Carlberg 2009; Smits et al. 2010) and 'limited negative results [-]' for the m-Trost (Smits et al. 2010). Due to the low sample size ($n < 25$), the evidence level of the test-retest reliability of the kinematic (Engsberg et al. 2004) and torque steadiness measurement (Bandholm et al. 2009) studies was scored as 'unknown [?]'

Table 9: Study 3 - General characteristics, psychometric properties and clinical utility SVMC measures

General Characteristic				Psychometric Properties		Clinical Utility				
Instrument	Body part tested	Scale	Outcome	Reliability (BoxB)*	Validity (BoxD,F,H)*	Clarity of instructions	Format	Qualifications	Time	Costs
b7600 (control of simple voluntary movements)										
Selective Motor Control	ankle	ordinal (0-4)	max 4 points leg score	Inter-rater: fair-good Loving and Carlberg (2009); Smits et al. (2010) Test-Retest: fair Loving and Carlberg (2009)	No study	excellent	assessment	not addressed	5 min	none
modified Trost test	ankle, knee, hip	ordinal (0-2)	max 8 points leg score	Inter-rater: fair Smits et al. (2010)	No study	excellent	assessment	recommended	15 min	none
Gillet's SMC test	for all joints	ordinal (0-2)	max 2 points per joint	No study	Construct (only for knee-flexion): Manikowska et al. (2016)	excellent	assessment	recommended	15 min	none
Selective Control Assessment of the Lower Extremity	toes, STJ, ankle, knee, hip	ordinal (0-2)	max 10 points leg score	Inter-rater: good-excellent Fowler et al. (2009); Balzer et al. (2016) Intra-rater: excellent Balzer et al. (2016)	Content: poor(Fowler et al. 2009) Construct: fair-good Fowler et al. (2009); Balzer et al. (2016); Kusumoto et al. (2016); Lim et al. (2015) Cohort: fair-good Fowler et al. (2010), (*)Fowler and Goldberg (2009); (*)Goldberg et al. (2011), Rha et al. (2015 and 2016) Criterion: excellent Balzer et al. (2016)	excellent	assessment	recommended	15 min	none
Kinematic	ankle	interval	end range ROM DE and PF SD relative phase measure (°)	Test-Retest: good Engsberg et al. (2004)	Cohort: fair Engsberg et al. (2004)	excellent	video recording system	required	not stated	high
EMG	swing phase knee	interval	hip-knee-angle-diagram (minimum relative phase)	No study	Cohort: fair Engsberg et al. (2008)	excellent	video recording system	required	not stated	high
				No study	(*)Fowler and Goldberg (2009)	adequate	3D gait analysis kinematic	required	not stated	high
b7300 (power of isolated muscles and muscle groups, muscle activation)										
Torque steadiness & EMG	ankle	interval	SD of target torques: DE: 0.1 and 0.3 Nm/kg body weight PF: 0.3 and 0.5 Nm/kg body weight	Test-Retest: good Bandholm et al. (2009)35	Cohort: fair Bandholm et al. (2009)35	excellent	customized equipment, isometric torque measure (plus EMG)	required	not stated	high
	hip, knee, ankle	interval	SD of [torque/mean torque] *100)	No study	Cohort: fair Arpin et al. (2013)36	excellent	isokinetic dynamometer (plus EMG)	required	not stated	high

ROM: range of motion; DE: Dorsal Flexion; PF: Plantar Flexion; SD: Standard Deviation; *COSMIN rating for every item (0-3: 0 = poor; 1 = fair; 2 = good; 3 = excellent) overall score for the methodological quality of the psychometric property assessed was determined by taking the lowest score of one box; ²⁸ Table is separated for ICF functions (b7600 and b7300).

Table 10: Study 3 - Summary study details, COMSIN-, quality and evidence-rating

ICF	Instrument	Study	Study Population		Construct validity: Hypothesis Testing (BoxF) (I)			Modified COSMIN**	Quality Criteria	Overall evidence
			Diagnose	Age (SD)	n	Design	Results	COSMIN*		
b7600 (control of simple voluntary movements)	Selective Control Assessment of the Lower Extremity (SCALE)	Fowler et al. (2009) ³⁷	spastic CP GMFCS I: 10; II:12; III: 19; IV:10	mean 11y, 11m (4y, 9m)	51	Construct validity: GMFCS	SCALE and GMFCS sig. rho: -0.83, sig	good	+	++ (SCALE – GMFCS)
		Fowler et al. (2010) ³⁸	spastic CP GMFCS I: 10; II: 10; III:18; IV:9	mean age 11y9m (4y8m)	48	Cohort study: pathophysiology (increased distal impairment, in relation to impaired corticospinal tracts)	- sig. differences in SCALE joint scores overall joint - sig. differences between all pairs of hip, knee, ankle - <i>proximal to distal concordance</i>	fair	+	++ (SCALE – joint)
		Fowler and Goldberg (2009) ⁴⁰	spastic CP GMFCS I-IV	range 6y-21y	15	Cohort study: SCALE and correlation inter-joint-coordination hip-knee angle diagrams during swing phase of gait	- higher SCALE scores normal hip-knee-angle-diagram; ROM increased - correlation total SCALE score and MRP: sig. rho 0.83 - correlation total SCALE and velocity: rho 0.66	poor	+	
		Goldberg et al. (2011) ³⁹	spastic CP GMFCS I:5; II:8; III:5	mean age 13y8m (7y2m)	18	Cohort study: correlation SCALE limb score and total swing joint contribution/swing extension and acceleration	- sig. PCC: 0.85 - as SCALE scores increased, the swing joint movements provided less resistance to knee extension - SCALE 0-4 = simultaneous hip and knee flexion = diminished knee extension acceleration	poor	+	++ (SCALE – knee swing phase)
		Rah et al. (2014) ⁴¹	spastic CP GMFCS I:5; II:8; III:5	mean age 10y 1m	34 (28)	Cohort study: regression: SCALE; contracture knee and ankle, gait-kinematics and muscle-tendon length	- SCALE scores and delayed timing of peak knee flexion during swing predicted sig. independently knee flexion at initial contact	poor	?	
		Baizer et al. (2016) ⁴²	spastic CP GMFCS I: 23; II:5; III:8; IV: 3	mean age 12y6m (3y7m)	39	Construct validity: GMFCS, limb distribution correlations: MAS MMT	- sig. differences between GMFCS levels (I vs II), most joint pairs, less and more affected leg - sig. correlation: MMT (rho: 0.88) MAS (rho: -0.55)	fair	+	++ (SCALE – GMFCS) (SCALE – joint)
		Lim et al. (2015) ⁴²	spastic CP GMFCS I: 5; II:8; III:7; IV: 3	mean age 9y3m (2y3m)	23	Construct validity: SCALE is a valid tool to predict the PBS	sig. differences between hemi- and diplegia, - sig. correlation for total and ltmes SCALE vs PBS (rho: 0.62-0.79)	poor	?	+
		Kusumoto et al. 2016) ⁴⁴	spastic CP GMFCS I: 12; II:19; III:9	mean 13y3m (3y4m)	40	Construct validity: correlation between SCALE and knee extensor strength	- sig. differences in SCALE scores more and less affected leg, no differences for knee extensor strength - inverse sig. correlation SCALE E and knee extensor strength (rho: 0.4)	fair	?	?

CP: cerebral palsy; GMFCS: Gross Motor Function Classification Level; rho: spearman's rank correlation coefficient; ROM: Range Of Motion; MRP: Minimum Relative Phase (measurement of interjoint coordination between the hip and the knee); PCC: Pearson product-moment correlation coefficient; MAS: Modified Ashworth Scale; MMT: Manual Muscle Testing; PBS: Pediatric Berg Balance Scale -COSMIN rating for every item (0-3: 0 = poor; 1 = fair; 2 = good; 3 = excellent) overall score for the methodological quality of the psychometric property assessed was determined by taking the lowest score of any items of one box.²⁸ ** modified COSMIN: omitting the sample size item boxB03; Quality Criteria: according to Terwee et al.²⁹ (Appendix 2): + = positive rating; ? = indeterminate rating; - = negative rating; Overall evidence: according to Cochrane van Tulder et al. 2003³⁰ (Appendix 3): strong (+++/+/-), moderate (++/-/-), limiting (+/-/-) = positive/ negative results; ± conflicting results; ? = unknown results.

Table 10. (continued)

ICF	Instrument	Study	Study Population			Construct validity: Hypothesis Testing (BoxF) (II)				Overall evidence
			Diagnose	Age (SD)	n	Design	Results	COSMIN*	Modified COSMIN**	Quality Criteria
b7600 (control of simple voluntary movements)	Gillett's SMC test	Mani-kowska et al. (2016) ³¹	CP GMFCS I-III	CP mean age 15y (5y5m); TD 22y (1y5m)	CP 23 TD 19	Observational study: CP muscle activity (EMG) during selective knee flexion versus neuro-logical intact controls	sig. differences between CP and Control, and between CP and different levels of SVMC (0 vs 2)	poor	poor	?
		Ergsberg et al. (2008) ³⁴	spastic CP bilateral GMFCS I: 5; II: 8; III: 1	CP mean age 16y (10y); TD age 15y (9y)	CP 29 TD 15	Cohort study: relative phase = quantification of the relative timing between a pair of oscillators at the same frequency	- sig. differences between CP (high SD) and TD (low SD) children in DE / PF movements - no sig. differences for antiphase movement	poor	fair	±
	Kinematic	Ergsberg et al. (2004) ³²	CP pre post SDR spastic CP	CP mean age 7y; TD mean age 7y2m	CP pre post SDR 12 CP 14 TD 20	Cohort study: differences between active ROM ankle in CP children pre and post SDR and vs TD	- sig. increase in active full ROM and active DE (but not PF) pre vs post SDR - no sig. differences in CP only group - sig. smaller ROM CP vs TD	poor	fair	(sample size)
		Zwaan et al. (2012) ³³	CP post SDR, (CP: GMFCS I-III)	CP mean age 6y5m; TD range 6-11y	CP post SDR 39 CP 38 TD 30	Cohort study: CP vs TD, correlation between synergies pattern (EMG), gait profile, mTrost, GMFM	- extensor synergy: sig. differences CP (0.95) vs TD (0.77) - thigh synergy: only sig. differences t for comfortable speed CP (0.94) vs TD (0.95) - no strong correlation between mTrost and gait EMG /GMFM	fair	fair	±
	EMG (thigh- and extensor-synergies)	Bandholm et al. (2009) ³⁵	CP unilateral, GMFCS I: 13, II: 1	CP mean age 11y (1y); TD mean age 11y (1y)	CP 14 TD 14	Cohort study: comparison CP and TD torque steadiness DE/PF	- sig. reduction of torque steadiness in CP (SD and CV were higher) vs TD - DE was most affected - sig. greater antagonist-agonist activation ratio and muscle activation variability in CP vs TD - DE torque steadiness correlated with level of PF coactivation (r = 0.597) and DE antagonist-agonist activation ratio (r = 0.832)	poor	fair	(sample size)
b7300 (power of isolated muscles and muscle groups, muscle activation)	Torque steadiness	Arpin et al. (2013) ³⁶	CP bilateral and unilateral, GMFCS I: 7; II: 5; III: 3	CP age mean 14y2m (7m); TD age mean 14y1m (7m)	CP 15 TD 15	Cohort study: comparison CP and TD torque steadiness DE/PF	- sig. greater CV at the ankle in CP vs TD - CP: sig. greater variability at ankle, then knee and hip - CP more regular steady torque patterns vs TD	poor	fair	(sample size)

CP: cerebral palsy; TD: Typically Developed children; GMFCS: Gross Motor Function Classification Level; SD: Standard Deviation; DE: Dorsal Extension; PF: Plantar Flexion; SDR: Selective Dorsal Rhizotomy; ROM: Range Of Motion; GMFM: Gross Motor Function Measure; CV: average coefficient of variation; rho: spearman's rank correlation coefficient; *COSMIN rating for every item (0-3: 0 = poor; 1 = fair; 2 = good; 3 = excellent) overall score for the methodological quality of the psychometric property assessed was determined by taking the lowest score of any items of one box; ** modified COSMIN: omitting the sample size item box B03; Quality Criteria: according to Terwee et al. ²⁹ (Appendix 2); + = positive rating; ? = indeterminate rating; - = negative rating; Overall evidence, according to Cochrane (van Tulder et al. 2003³⁰, Appendix 3): strong (+++/-), moderate (++/-), limiting (+/-) = positive/negative results; ±: conflicting results; ? = unknown results; Table is separated for ICF functions (b7600 and b7300).

Table 10. (continued)

ICF	Instrument	Study	Study Population			Content Validity (BoxD) / Criterion Validity (BoxH)				Overall evidence
			Diagnose	Age (SD)	n	Design	Results	COSMIN*	Modified COSMIN**	Quality Criteria
b7600 (control of simple voluntary movements)	Selective Control Assessment of the Lower Extremity (SCALE)	Fowler et al. (2009) ³⁷	spastic CP, GMFCS I: 10; II:12; III: 19; IV:10	mean 11y, 11m (4y, 9m)	51	Content validity: 14 experiences clinicians	mean agreement: 91.9% (range 71.4–100%) for content, administration, and grading	poor	poor	+
		Balzer et al. (2016) ⁴⁵	spastic CP GMFCS I: 23; II:5; III:8; IV: 3	mean age 12y6m (3y7m)	39	Criterion validity: correlation SCALE and FMA item III & IV	sig. correlation: FMA (rho 0.88)	fair	excellent	+

CP: cerebral palsy; GMFCS: Gross Motor Function Classification Level; rho: spearman's rank correlation coefficient; FMA: Fugl-Meyer Assessment; *COSMIN rating for every item (0-3; 0 = poor; 1 = fair; 2 = good; 3 = excellent) overall score for the methodological quality of the psychometric property assessed was determined by taking the lowest score of any items of one box; ²⁸ ** modified COSMIN: omitting the sample size item box B03; *Quality Criteria:* according to Tenwee et al. ²⁹ (Appendix 2): + = positive rating; ? = indeterminate rating; - = negative rating; Overall evidence: according to Cochrane (van Tulder et al. 2003³⁰, Appendix 3): strong (+++/+/-/-), moderate (++/-/-/-), limiting (+/-/-) = positive/ negative results; ± conflicting results; ? = unknown results.

Table 10 (continued)

Study Population				Reliability (Box B)								
ICF	Instrument	Study	Diagnose	Age (SD)	n	Design	Results	COSMIN*	Modified COSMIN**	Quality Criteria	Overall evidence	
b7600 (control of simple voluntary movements)	Selective Motor Control test (SMC)	Lowling and Carlberg (2009) ⁴⁵	CP bilateral n=30; unilateral n=10; GMFCS I:13; II:12; III:10; IV:3 and V:2	median age 7y; range 3y-16y	29	Test-Retest (1-60 days)	Kw ankle DE = 0.88-1; RV: 0 - 0.005; RP: -0.11 - 0; RC: 0 - 0.05 (right & left leg scores)	poor	fair	+	+	(SMC Test-Reltest)
					40	Interrater (3 PTs)	Kw ankle DE: 0.58-0.77 RV: 0.000-0.030; RP: -0.09 - 0.11; RC: -0.26 - 1.16 (right & left leg scores)	fair	good	-	- - (SMC interrater)	
	Modified Trost (m-Trost)	Smits et al. (2010) ⁴⁶	spastic CP	mean age 6y,5m (12m)	21	Interrater (1 PT and 1 Dr) 1h in-between SMC and mTrost	K ankle DE: 0.55 95% IC: 0.36-0.74 (leg total score) K ankle DE: 0.65; 95% IC: 0.47-0.84 K knee EXT: 0.69; 95% IC: 0.49-0.88 K hip ABD: 0.57; 95% IC: 0.37-0.78 K hip FLEX: 0.71; 95% IC: 0.51-0.91	poor	fair	-	- (m-Trost interrater)	
					20	Interrater (2 groups: 3 PT and 3 Dr)	ICC: 0.88-0.91; 95% CI: 0.69-0.97 (right & left leg scores)	poor	good	+	+	(SCALE interrater)
	b7300 (power of isolated muscles and muscle groups, muscle activation)	Selective Control Assessment of the Lower Extremity (SCALE)	Fowler et al. (2009) ³⁷	spastic CP	mean age 12y,3m (5y5m)	38	Interrater (2 PTs via Video) Intrater (1 rater via video, 6 weeks)	ICC:0.91-0.94 (less & more affected leg); MMD: 1.88-1.92 ICC:0.95-0.96 (less & more affected leg); MMD: 1.79-1.96	fair	excellent	+	+
8						Test-Retest (8 weeks)	sig. PCC: ankle PF: 0.77; ankle DE: 0.94; total sagittal range: 0.93 mean differences of: max PF: 1.6°; DE = 1.1°; total range = 2.4°	poor	good	+	?	(sample size) (Test-Reltest)
Torque steadiness		Bandholm et al. (2009) ³⁵	CP	not stated (pilot-study)	7	Test-Retest (1 day)	ICC ankle PF:0.72; CV < 19% ICC ankle DE: 0.31; CV 25%	poor	good	+	?	(sample size) (Test-Reltest)
CP: cerebral palsy; GMFCS: Gross Motor Function Classification Level; Kw: Kappa weighted; DE: Dorsal Flexion; PF: Plantar Flexion; RV: Relative rank Variance (random disagreement); RP: Relative Position; RC: Relative Concentration (systematic disagreement); PT: Physiotherapist; Dr: Physician; K: Kappa; 95% IC: 95% Confidence interval; EXT: Extension; ABD: Abduction; FLEX: Flexion; ICC: Intraclass Correlation Coefficient; MMD: Minimal detectable change; PCC: Pearson product-moment correlation coefficient; CV: average coefficient of variation, *COSMIN rating for every item (0-3: 0 = poor; 1 = fair; 2 = good; 3 = excellent) overall score for the methodological quality of the psychometric property assessed was determined by taking the lowest score of any items of one box; ²⁸ ** modified COSMIN: omitting the sample size item box B03; <i>Quality Criteria</i> : according to Terwee et al. ²⁸ (Appendix 2); + = positive rating; ? = indeterminate rating; - = negative rating; Overall evidence: according to Cochrane (van Tulder et al. 2003 ³⁰ , Appendix 3): strong (+++/+/-), limiting (+/-) = positive/ negative results; ±: conflicting results; ? = unknown results; Table is separated for ICF functions (b7600 and b7300).												

CP: cerebral palsy; GMFCS: Gross Motor Function Classification Level; Kw: Kappa weighted; DE: Dorsal Flexion; PF: Plantar Flexion; RV: Relative rank Variance (random disagreement); RP: Relative Position; RC: Relative Concentration (systematic disagreement); PT: Physiotherapist; Dr: Physician; K: Kappa; 95% IC: 95% Confidence Interval; EXT: Extension; ABD: Abduction; FLEX: Flexion; ICC: Intraclass Correlation Coefficient; MMD: Minimal detectable change; PCC: Pearson product-moment correlation coefficient; CV: average coefficient of variation; *COSMIN rating for every item (0-3: 0 = poor; 1 = fair; 2 = good; 3 = excellent) overall score for the methodological quality of the psychometric property assessed was determined by taking the lowest score of any items of one box; ** modified COSMIN: omitting the sample size item box B03; Quality Criteria: according to Tenwee et al. (Appendix 2); + = positive rating; ? = indeterminate rating; - = negative rating; Overall evidence: according to Cochrane (van Tulder et al. 2003³⁰; Appendix 3); strong (+++/+/-), moderate (++/-/-), limiting (+/-/-) = positive/ negative results; ± conflicting results; ? = unknown results; Table is separated for ICF functions (b7600 and b7300).

5.5 Discussion

This review revealed a limited number of psychometric studies investigating SVMC measures in children with UMN lesions. The overall evidence was further limited, because 10 out of 17 studies were cohort studies with a limited methodological quality (e.g. 'poor' or 'fair') - except one study (Fowler and Goldberg 2009), which scored 'good' according to modified COSMIN rating guidelines. No study investigated responsiveness, which would be crucial regarding the measurement of therapy-induced changes of SVMC.

The chosen age range (2-21 years) for this review, might have been wide when considering developmental issues which are known to influence SVMC (e.g. maturation of CNS function), as well as the importance of the participants' cognitive understanding and motivation for the SVMC measurement procedure/testing. Nevertheless, this age range was chosen due to the overall limited number of studies available for review. Future studies regarding SVMC measures may choose to investigate psychometric properties in separate age groups (e.g. pre CNS maturation 2y-7y and post < 8 years).

5.5.1 SVMC Assessment Tools

The SCALE was the most often investigated assessment tool, in terms of the number of studies conducted and in the number of its measurement properties investigated. The SMC and m-Trost were only rated on reliability, thus lacking evidence on their validity. The Gillette's SMC test was only investigated on its validity, lacking evidence about its reliability. In terms of psychometric quality, the SCALE had the highest level of evidence with a moderate positive level of evidence concerning its inter-rater reliability and its construct validity, and an unknown and limited level of evidence of content and criterion validity, respectively. Another advantage of the SCALE in comparison to the other assessment tools (SMC, m-Trost,) lies in its evaluation of five lower extremity joints rather than one or three joints. In addition, clinical utility (Table 9) of the SCALE, as well as of the other SMC assessment tools, was scored high as time, costs and resources are low. However, in terms of limitations, the SCALE's ordinal scoring system relies on the impression of the rater (e.g. therapist, consultant), which make it to a subjective measurement. Finally, as discussed already in the last chapter, the SCALE's ordinal scoring system (normal, impaired, and unable) may lack sensitivity to detect certain therapy-induced changes of SVMC.

5.5.2 SVMC laboratory-based Measures

The construct validity of the kinematic, sEMG and torque steadiness was assessed, but none of the papers evaluating these measurement techniques explicitly mentioned that the assessment of validity was an a priori objective. Because of this, the formulation of hypotheses was often absent thus diminishing their modified COSMIN score to 'fair'. Only two laboratory-based SVMC measures (kinematic (Engsberg et al. 2004) and torque steadiness (Bandholm et al. 2009)) were assessed regarding their test-retest reliability. In terms of psychometric quality as well as clinical utility (see Table 9), none of the identified laboratory-based measures seem to offer a great advantage over the other. The equipment that was required to record the outcome measures was often customized, making it difficult for other groups (researchers or clinicians) to apply and confirm or extend findings of studies exploring the laboratory-based measures using sEMG, kinematics or torque measurements. Furthermore, the measurement procedures appear time consuming and complex in comparison to more routinely applied clinical assessments. Personnel also required extensive training in the application and analysis of these measures (see Table 9).

In summary, the results from this systematic review show the limited level of evidence regarding the psychometric properties (reliability and validity) and absence of evidence regarding the responsiveness of currently available SVMC measures of the lower extremity in children with UMN lesions.

5.5.3 Limitations and methodological considerations

Low inter-rater agreement when rating the quality of the evidence in systematic reviews (e.g. rating Risk of Bias in Cochrane type reviews) can be an important methodological issue, which should be considered when conducting a systematic review (Jørgensen et al. 2016). In this review, agreement between the raters for all COSMIN items was high. Only five out of 246 items needed further discussion, and none required the rating of a third reviewer. This high agreement was likely the result of the specific rating rules, which we established as recommended by the COSMIN group. For example, when scoring the reliability items 4-7 for the SCALE, we decided in advance to score the use of video for the evaluating of the inter- and intra-rater reliability as appropriate, because this allows a discrete evaluation of the scoring system by maintaining the stability of test conditions and patient status, as well as saving on time and resources. In contrast, a video approach was not considered to

be appropriate for determining test-retest reliability when the stability of the patient is evaluated.

In line with other neuro-paediatric COSMIN reviews (Ammann-Reiffer et al. 2014; Gerber et al. 2016; Lennon et al. 2015), the rating of the sample size item (modified COSMIN score) was modified. This is, because the sample size is often limited in clinical neuro-paediatric studies and not comparable with large scale epidemiological healthcare studies using patient-reported outcome measures for which the COSMIN guidelines were initially evaluated. This modified scoring improved the overall rating of all studies with the exception of the construct and content validity score of the studies from Fowler et al. (2009) and Zwaan et al. (2012). Although this modified score was, it would be recommended for future psychometric studies to include studies with a sufficiently large sample size (>30).

Other reasons for scoring poor were the lack of 'a priori' formulated hypotheses (box 'hypothesis testing') and for one study (Fowler et al. 2009) the lack of evaluating each item separately for its content validity (box 'content validity'). While it was considered to be important that each single question should be evaluated separately for its content validity in a Health Care Questionnaire (where the COSMIN was originally developed for), it could be questioned whether the same rating rules are necessary for an assessment tool like the SCALE that consists of a similar procedure repeated for different joints.

5.5.4 Clinical implications and future work

The results of this review show that the SCALE is the most frequently investigated assessment method in the population of CP children and also, that it has the highest quality rating of the evidence of its psychometric properties. Its responsiveness to change has not been assessed, but it may be expected that due to its ordinal scoring system its sensitivity to measure changes of SVMC is limited. To improve its sensitivity and simultaneously to benefit from its child-friendly procedure, combining the SCALE with another, more sensitive measure appears to be promising. This idea has also been proposed in the previous study of this PhD project (Chapter 4, Balzer et al. 2016) as well as in other studies (Zwaan et al. 2012; Dobson 2010). While Zwaan et al. (2012) found no convincing evidence for detecting extensor and flexor synergies during gait using sEMG in children with CP, they reported a significant cross-correlation between extensor synergy activities measured using sEMG and the m-

Trost test. They concluded that “sEMG measures still may be useful for selective motor control measurement because it measures selectivity at the level of the specific muscles involved, provided the appropriate task is used” (Zwaan et al. 2012). As walking requires selective as well as synergistic movements, this may not be an appropriate task for the assessment of SVMC. The tasks embedded in the SCALE (isolated single-joint movements) were developed in accordance with the definition of SVMC (Sanger et al. 2006; Fowler et al. 2009). Combining SCALE’s ratings for single-joint movements with sEMG, recordings would further allow for directly measuring voluntary activation of a muscle even in patients with low muscle strength (manual muscle test grade of 1), whereby no real joint movement occurs.

5.6 Conclusion

This systematic review revealed a limited number of psychometric studies evaluating the validity and reliability of SVMC measures in children with UMN lesions, whereas no studies evaluated responsiveness. The SCALE appeared to have the highest level of evidence regarding its reliability and construct validity compared to other clinical and laboratory-based measures of SVMC. Nevertheless, SCALE’s application as a measure to assess therapy-induced changes in SVMC is questionable due to its broad scoring system. Concrete ideas on how to expand SCALE’s application to sensitively measure therapy-induced changes of SVMC in children with UMN lesions were present.

Chapter 6: Study 4

6.1 Purpose of chapter

This chapter provides an overview of how the last study developed within the context of this PhD project. Furthermore, research questions 10 to 13 of Section 2.4 will be answered by describing the chosen method and presenting as well as discussing the study results.

This chapter is in preparation for being published in the Journal Neurorehabilitation and Neural Repair.

6.2 Background

The results of the previously described three studies (Balzer et al. 2016; Balzer, Marsico, et al. 2017; Balzer, van der Linden, et al. 2017) led to the design of this fourth study. First, the regression study showed the relevance of SVMC for mobility in children with CP and thereby underlined the importance of determining this negative UMN sign (Balzer, Marsico, et al. 2017). Second, the appropriate feasibility, validity and reliability of the SCALE as a diagnostic and prognostic assessment tool was confirmed. Third, the systemic review (Balzer, van der Linden, et al. 2017) revealed a lack of responsive SVMC outcome measures that are required detecting therapy-induced changes of SVMC. In addition, the results of the systematic review also showed that the SCALE was the tool with the highest level of evidence regarding its reliability, validity and clinical utility. Consequently, the idea arose to use the SCALE testing procedure, and to measure the level of SVMC via an appropriate interval-scaled measurement tool. In accordance with the systematic review results, EMG is suggested as a suitable measure to collect during the SCALE testing due to the following reasons: i) it measures SVMC in accordance to the definition of SVMC: focusing on muscle activation during single-joint movement (Sanger et al. 2006), ii) it can detect voluntary activations without actual or during small joint movements (allows for measuring patients with muscle weakness as well); iii) it is also applicable for bilateral involvement (detecting mirror movements) and iv) its child friendly application (e.g. short duration, non-invasive). However, besides these benefits of sEMG, its limitations when used as a repeated measure are also well known. Differences in electrode placement, muscle size, subcutaneous fat thickness and muscle-fibre geometry can lead to absolute amplitude differences both within and

between participants. Normalization techniques to minimize this variability have been shown to be less appropriate in children with CP, due to their inability to reliably perform maximal voluntary contractions (Shuman et al. 2017; Zwaan et al. 2012; Bojanic et al. 2011). The Similarity Index (SI) is an outcome measure derived from sEMG, which uses bilateral prototype activation-patterns (collected from a neurologically healthy control group) in order to decrease the impact of this inter and intra subject variability. In the original paper (Lee et al. 2004), the SI is calculated as part of the Voluntary Response Index (VRI), that was developed for measuring selectivity in patients with spinal cord injury. The SI was one of the outcome measures resulting from the initial systematic search (study 3) aiming to identify SVMC tools used in the entire population of UMN lesions (Balzer, van der Linden, et al. 2017). The SI is a measure of the similarity between the sEMG pattern of lower leg muscles (5 for each leg) during a maximal standardized movement derived from a participant with a UMN and average pattern derived from a health control group. Mathematically, the participants' SVMC is thereby expressed as a ten-dimensional vector, which is compared in length and direction with the norm-vector. The SI can take any value between 1 to 0. While values close to "1" indicate that the SVMC is (near) normal, lower values indicate that the sEMG pattern of the movement deviates from the norm. Because the SI is derived from both legs, co- and mirror movements lower the value of SI, hence indicating a more abnormal sEMG pattern. However, the SI takes every muscle activation into account and cannot differentiate between physiological prime mover activities and pathological co- and/or mirror activities. As such, the SI is an unweighted measure, which compares the overall activation pattern of all involved muscle with the mean activation pattern recorded for the healthy population (norm).

In summary, the sEMG based SI algorithm has the following properties. I) It employs a normalization technique by including norm data from a neurologically intact control group. Thereby, it is not dependent on other normalization techniques (i.e. maximal voluntary contractions), which have shown to be inaccurate when performed with children with CP. II) The SI can detect, although not differentiate bilateral co- and mirror-activations, which are important in relation to the definition of SVMC. III) It uses a standardized testing procedure (single and multi-joint movements), which is similar to the testing movements of the SCALE procedure. Therefore, calculating the SI algorithm when using the SCALE procedure for testing seems feasible. IV) It has acceptable validity and reliability in patients with spinal cord injuries (McKay et al. 2005; Lim and Sherwood 2005; Lee et al. 2004). V) It measures SVMC on an interval

scale. Accordingly, the SI seemed to be an appropriate interval measure, capable of measuring SVMC while asking the child to perform the SCALE testing procedure.

Therefore, the aim of this study was to investigate the validity and reliability of the 'SCALE-SI', in which the SI algorithm objectively quantified SVMC while the child is performing the SCALE testing procedure. In line with study 2 (Balzer et al. 2016) as well as with the assessment of the validity of the SCALE's performed by Fowler et al. (2009), the following hypotheses were formulated to investigate the validity of the SCALE-SI. Firstly, to establish concurrent validity of the SCALE-SI, a high positive correlation ($p > 0.70$) between the SCALE-SI and the SCALE was expected. Secondly, concerning its discriminative validity, the following statistically significant differences were expected: i) between children with CP and neurological intact children; ii) across all and between SCALE-SI joint scores (hip vs. knee) for each limb (Children with CP only), and iii) between the less and more affected limb (children with CP only). Finally, test-retest reliability should be at least moderate, with ICC values exceeding 0.65 (similar levels as in the study of Lim and Sherwood 2005) and acceptable absolute measurement errors.

6.3 Method

6.3.1 Participants

Thirty-one neurologically intact adults were recruited by convenience sampling in order to collect norm values needed for the calculation of the SI. Adults between 18 and 50 years, who were symptom-free in terms of any central or peripheral neurological injury and who had no surgery of the lower limbs within the last year, were included.

Furthermore, in- and out-patients of the "Rehabilitation Centre Affoltern am Albis, University Children's Hospital Zurich" were recruited by convenience sampling from June 2017 until March 2018. According to recommendations for adequate sample size for assessment of validity, we aimed to include 30 children with CP (Mokkink and Terwee 2010). Inclusion criteria were: clinical diagnosis of spastic or mixed CP, age between 5 and 20 years, and ability to follow simple instructions. Children with a primarily dystonic or ataxic impairment, an unstable situation regarding their tonus-regulating medications and/or children who had a botulinum toxin injection within the last 6 months or any surgical correction within the last year of the lower extremity were excluded. For establishing discriminative validity, 32 neurologically intact children

were recruited. Only children and young people aged 6-18 years, and without any medical history of neurological and/or orthopaedic diagnosis within the lower extremity were included.

The study was approved by the ethical committee of the Canton of Zurich (KEK-ZH-Nr.2011-0404). Informed consent and assent were obtained from parents and participants, respectively.

6.3.2 Measurements

All tests were carried out by a team of three testers (one experienced neuro-paediatric physiotherapist and two human movement scientists) within a maximum timeframe of 1h and in accordance to standardized procedures.

In order to gain information about the degree of spasticity, the MAS (0-4) was applied for flexors and extensors muscle of the hip, knee and ankle joint (Bohannon and Smith 1987).

For the SCALE-SI, which consisted of the SCALE testing procedure while simultaneously recording bilateral multichannel sEMG activities, the following sEMG setting was applied. Self-adhesive Ag/AgCl dual snap gel electrodes (Noraxon Inc, Scottsdale/USA) with a diameter of 10 mm and an inter-electrode distance of 20 mm were applied bilaterally in accordance to SENIAM guidelines to the following muscles: m. gastrocnemius medialis (GM), m. peroneus longus (PL), m. tibialis anterior (TA), m. rectus femoris (RF), and m. semitendinosus (ST). The recorded sEMG data was then offline processed with the SI algorithm to quantify SVMC. The participant was then positioned on a custom-made wooden seat-rest, which had been furnished with special openings for the sEMG sensors of the ST (Figure 11a).

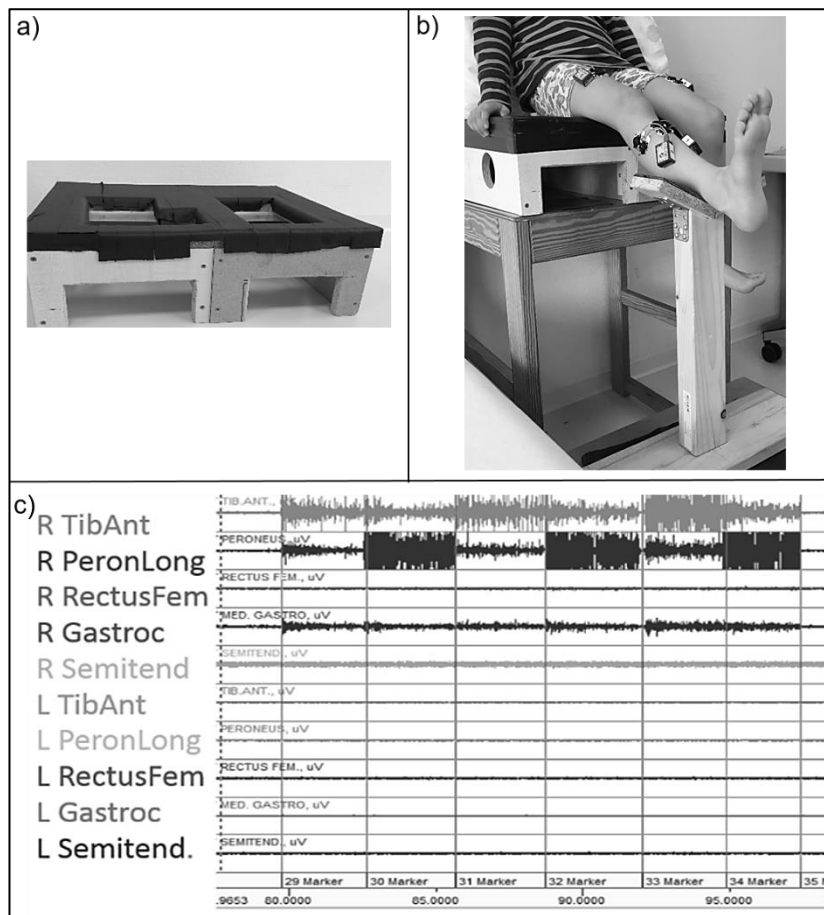


Figure 11: Study 4 - SCALE-SI testing setup:

a) Customized testing seat for avoiding pressure on the sEMG sensors of the m. semitendinosus; b) Leg-holder for ankle and STJ movements; c) Recorded norm SCALE-SI sEMG-patterns of the right Subtalar Joint (STJ) data set STJ movements, Note: Trigger-Markers for each movement direction (event-window).

Abbreviations: R / L TibAnt: right / left m. tibialis anterior; R / L PeronLong: right / left m. peroneus longus; R / L RectusFem: right / left m. rectus femoris, R / L Gastroc: right / left m. gastrocnemius; R / L Semitend: right / left m. semitendinosus

First, baseline activity of one minute was recorded in sitting position, while the legs were hanging relaxed. Then, the original SCALE testing procedure was carried out to assess SVMC at the hip, knee, ankle and subtalar joints. SVMC of the toes was not measured due to impracticability of applying EMG on small foot-muscles of children. In order to minimize positional changes, which are known to interfere with sEMG measures, the SCALE procedure was slightly adjusted by, firstly, testing SVMC of the hip in sitting (instead of lying) position and, secondly, by changing the order of testing of the joints: knee, ankle, STJ, hip (instead of hip, knee, ankle, STJ). Furthermore, to maximize relaxation of the muscles, which would not become activated under physiological movement, the following adaptations were made, and the following standardized instructions were given. The lower leg was positioned on a pedestal just

proximal to the ankle joint to allow for maximal relaxation of the RF while assessing ankle and STJ movements (Figure 11.B). Before each joint movement, the child was instructed to relax all other muscles as much as possible and to focus just on the muscles, which were used to move the body segment.

In line with determining leg dominance (Chapman et al. 1987), we determined the less affected leg by asking participants with which leg they kick a ball. In children with CP, we also relied on the diagnosis (e.g. in cases of unilateral spastic CP).

For the assessment of test-retest reliability of the SCALE-SI in children with CP, the measurement was repeated by the same testing team under similar conditions (time of day, room). In order to have stable yet independent measurements, assessments were repeated seven to 15 days later and data were written on a new blank collection form.

6.3.3 Data recording and processing

Data acquisition was performed with a 16-channel Myosystem 1400A (Noraxon Inc, Scottsdale/USA) with a sample frequency of 1000 Hz and a bandwidth of 30–500 Hz. sEMG patterns of the ten muscles were recorded during the four joint movements within one session. In line with the SCALE manual, each joint movement was performed bidirectionally (e.g. knee: extension, flexion) and fluently repeated three times (e.g. no rest, in one smooth action). Event markers, identifying the start and end of each of these voluntary movements (e.g. knee: extension, flexion, extension, flexion, extension, flexion) (Figure 11c), were triggered manually by the tester who observed the movements. A total of 56 event markers was set per SCALE-SI procedure (e.g. per joint, seven markers to identify the three repetitions of each movement direction (flexion / extension)); this was repeated for each of the eight joints that were tested).

All sEMG signals were filtered using a 20 Hz high pass filter (Finite Impulse Response (FIR) filter). An Infinite Impulse Response (IIR), 50Hz Rejection filter, was only used for signals which showed 50Hz noise due to other electronic devices. Correct event marker placement as well as artefact detection took place via visual off-line inspection. Position and reason for the all movement-artefacts were documented for the further data processing. Movement artefacts occurred either due to involuntary sEMG-cable contact with seat or due to manual sEMG-sensor contact by the participants.

Further sEMG data analysis was performed using Matlab (version 2017b). Firstly, movement artefacts were corrected manually by removing (minimal cutting) them from the data. Second, baseline correction was performed by using the lowest activity values for each of the 10 muscles recorded during the test trial itself. This deviation from the original protocol was necessary, because the muscle activity during the initial baseline-trial was in many participants higher than during the SCALE-SI trial, causing negative values. Next, the sEMG data within each movement event-window was rectified using a root mean square (RMS) algorithm. The rectified individual manoeuvre sEMG data were then averaged across the three repetitions of each motor task (e.g. extension-flexion, extension-flexion, and extension-flexion). The SI quantifies the similarity between the sEMG activation-pattern of the participant and the norm (derived from the adult control group) and was calculated using the algorithm described by Aslan et al. (2013). The SI is one of the numeric outcomes of the VRI, which furthermore comprises of the “Magnitude” (Mag). The Mag is derived from the total mean activity of all ten muscles recorded during the voluntary motor task. It represents the length of the ten-dimensional vector. The following four equations were used:

Equation 1: Response Vector (RV): The RV consists of the mean activity of each of the ten muscles, averaged over the three movement repetitions (e.g. extension-flexion), and was derived for each of the eight joint movements per leg. Quantitatively, the RV describes the absolute activity in each muscle during each phase of each movement.

$$RV = \begin{bmatrix} r_{TA} \\ r_{PL} \\ r_{RF} \\ r_{GM} \\ r_{ST} \\ l_{TA} \\ l_{PL} \\ l_{RF} \\ l_{GM} \\ l_{ST} \end{bmatrix}$$

(Aslan et al. 2013)

Equation 2: Magnitude (Mag): The magnitude of the RV is calculated by using the square root of the sum of the squares of the activity of each of the 10 muscles; it represents the length of this vector.

$$\text{Magnitude RV participant} = |RV| = \sqrt{(r_{TA}^2 + r_{PL}^2 + r_{RF}^2 + r_{GM}^2 + r_{ST}^2 + l_{TA}^2 + l_{PL}^2 + l_{RF}^2 + l_{GM}^2 + l_{ST}^2)}$$

(Aslan et al. 2013)

Equation 3: The Prototype Response Vector (PRV) was computed by averaging the RVs of the 31 neurologically intact adults.

$$PRV = \begin{pmatrix} PRVrTA \\ PRVrPL \\ PRVrRF \\ PRVrGM \\ PRVrST \\ PRVITA \\ PRVIPL \\ PRVIRF \\ PRVIGM \\ PRVIST \end{pmatrix} = \begin{pmatrix} (RrTA1 + RrTA2 + RrTA3 + \dots + RrTA31) / N \\ (RrPL1 + RrPL2 + RrPL3 + \dots + RrPL31) / N \\ (RrRF1 + RrRF2 + RrRF3 + \dots + RrRF31) / N \\ (RrGV1 + RrGV2 + RrGV3 + \dots + RrGV31) / N \\ (RrST1 + RrST2 + RrST3 + \dots + RrST31) / N \\ (RITA1 + RITA2 + RITA3 + \dots + RITA31) / N \\ (RIPL1 + RIPL2 + RIPL3 + \dots + RIPL31) / N \\ (RIRF1 + RIRF2 + RIRF3 + \dots + RIRF31) / N \\ (RIGM1 + RIGM2 + RIGM3 + \dots + RIGV31) / N \\ (RIST1 + RIST2 + RIST3 + \dots + RIST31) / N \end{pmatrix}$$

(Aslan et al. 2013)

Equation 4: Similarity Index (SI): The SI is the scalar product of the norm response vector and a participant's response vector divided by the product of their magnitudes. It ranges from 0 to 1. Values close to "1" indicate that the participants' SVMC is very similar to that of the norm.

$$SI = \frac{\sum_i (PRV_i * RV_i)}{|PRV| * |RV|}$$

(Aslan et al. 2013)

6.3.4 Statistical Analysis

As the Shapiro-Wilk-tests showed that the majority of SCALE-SI scores (joint, leg, total) were not normally distributed, non-parametric tests were applied to test our a priori formulated research hypotheses. Data were presented for the less and more affected leg.

Concurrent validity was evaluated by calculating Spearman's rank correlation coefficients ($\rho = 1 - (6\sum d^2) / n(n^2 - 1)$) between SCALE-SI and conventional SCALE scores for each separate joint (joint score: 2-0), limb (summed joint score per leg: 0-10), and the total scores (summed leg score: 0-20), respectively, as well as between the SCALE-SI total score and the GMFCS level.

The discriminative validity of the SCALS-SI was determined using the following statistical tests: i) a Mann-Whitney U Test to compare differences of the SCALE-SI scores (separate joints, limb, total) between children with CP and their neurologically intact peers; ii) a Friedman Test to determine whether SCALE-SI scores differed between the four adjacent joint pairs within a limb and consecutive post-hoc Page tests in order to evaluate differences between joint pairs (e.g., hip versus knee), and iii) a Wilcoxon Signed Rank Test to investigate whether differences exist between the

less and more affected leg within a group of children with either bilateral or unilateral leg involvement. Effect size (r) was calculated in accordance to Rosenthal (1994) ($r = |Z / \sqrt{N}|$). Alpha was set at 0.05 (two-tailed) (Portney and Watkins 2000).

Relative test-retest reliability was evaluated by a two-way random ANOVA model ($ICC(2,1) = MSR - MSE / MSR + (k-1) MSE + (k/n)(MSC - MSE)$ (Koo and Li 2016)), as each participant was rated by a single rater and the results were to be generalized to other raters. Correspondingly, 95% confidence intervals were calculated for joint and leg scores as well.

Absolute reliability was determined by the SEM ($SEM = SD \text{ (of test and retest)} \times \sqrt{1 - r_{xx}} \text{ (ICC2.1)}$) and the MDC ($MDC = SEM \times \sqrt{2} \times 1.96$) (Haley and Fragala-Pinkham 2006). Statistical analysis was performed with R. Studio (version 3.4.0).

6.4 Results

Twenty-four children with spastic and mixed type of CP (bilateral $n=20$; unilateral $n=4$) gave informed consent to participate in this study. Age ranged from 6y9m to 17y4m. Fourteen participants were male. Nine children had GMFCS level I, four had level II, five level III and six level IV. Further characteristics are presented in table 11. Due to organizational issues, three children could not participate in reliability testing. Therefore, 24 data sets were available for validity testing and 21 for reliability testing.

Table 11: Study 4 – Participants' characteristics

Participants' characteristics: children with CP (n=24)											
	less affected leg					more affected leg					Total
	hip	knee	ankle	STJ	leg	hip	knee	ankle	STJ	leg	
SCALE-SI (n=24)											
mean	0.75	0.76	0.84	0.83	0.79	0.71	0.72	0.72	0.67	0.70	0.75
Sd	0.17	0.18	0.19	0.15	0.14	0.19	0.20	0.26	0.24	0.17	0.14
median	0.76	0.78	0.93	0.89	0.83	0.77	0.74	0.80	0.63	0.71	0.75
IQR	0.30	0.29	0.20	0.26	0.17	0.19	0.36	0.39	0.35	0.19	0.26
range	0.39-0.98	0.40-0.97	0.37-0.99	0.52-0.97	0.50-0.94	0.31-0.99	0.39-0.99	0.19-0.99	0.16-0.99	0.41-0.95	0.19-0.52
SCALE (n=24)											
mean	1.38	1.33	1.38	1.12	5.21	1.25	1.21	0.92	0.62	4	9.21
Sd	0.58	0.64	0.71	0.9	2.26	0.53	0.59	0.83	0.65	2.11	4.04
median	1	1	1.5	1	5.5	1	1	1	1	4	9.5
IQR	1	1	1	2	3.25	1	1	2	1	4	0
range	0-2	0-2	0-2	0-2	1-8	0-7	0-2	0-2	0-2	0-2	1-15
MAS*(n=21)											
mean	0.55	0.95	0.85	NT	2.63	0.60	1.10	1.13	NT	3.19	5.81
Sd	1.14	1.06	0.75	NT	2.53	1.06	1.32	0.77	NT	2.76	5.06
median	0	1	1	NT	2.00	0	1	1	NT	2.00	4.00
IQR	1	2	1	NT	2.25	1	2	1	NT	4.25	4.75
Range	0-1	0-3	0-2	NT	0-8	0-2	0-3	0-2	NT	0-9	0-17

Abbreviations: SCALE-SI: Selective Control Assessment of the Lower Extremity – Similarity Index; MAS: Modified Ashworth Scale; sd: standard deviation; IQR: Interquartile Range. MAS* (n=21): Due to organizational issues, three children could not participate in reliability testing.

A descriptive summary of the norm (neurological intact adults n=31 (female n=14); mean age of 33y 7mo [SD 7y 4mo]) and control group (neurological intact children n=32 (female n=16); age range from 6y to 17y6m [SD 3y 5mo]) is shown below (Table 12).

Table 12: Study 4 – Norm and control group characteristics

Norm group: neurologically intact adults (n=31)											
	less affected leg					more affected leg					Total
	hip	knee	ankle	STJ	leg	hip	knee	ankle	STJ	leg	
SCALE-SI											
mean	0.93	0.93	0.96	0.94	0.94	0.94	0.95	0.95	0.91	0.94	0.94
sd	0.10	0.10	0.07	0.07	0.04	0.09	0.04	0.10	0.09	0.05	0.04
median	0.98	0.97	0.99	0.95	0.95	0.97	0.96	0.98	0.94	0.95	0.95
IQR	0.06	0.05	0.04	0.06	0.00	0.05	0.03	0.03	0.09	0.00	0.00
range	0.54-0.99	0.57-0.99	0.61-1	0.66-1	0.85-0.99	0.66-0.99	0.84-0.99	0.45-1	0.64-1	0.76-0.99	0.83-0.98
Control group: neurologically intact children (n=32)											
	less affected leg					more affected leg					Total
	hip	knee	ankle	STJ	leg	hip	knee	ankle	STJ	leg	
SCALE-SI											
mean	0.91	0.89	0.95	0.92	0.92	0.93	0.88	0.92	0.93	0.91	0.92
sd	0.13	0.13	0.06	0.08	0.06	0.10	0.14	0.09	0.08	0.06	0.05
median	0.98	0.95	0.97	0.94	0.94	0.97	0.94	0.97	0.95	0.93	0.93
IQR	0.10	0.12	0.06	0.05	0.00	0.11	0.10	0.11	0.05	0.00	0.00
range	0.47-0.99	0.35-0.99	0.78 - 1	0.66-1	0.73-0.98	0.56-0.99	0.33-0.99	0.58-1	0.60-1	0.73-0.89	0.75-0.98
SCALE											
mean	1.81	1.61	1.71	1.71	8.35	1.90	1.77	1.81	1.94	7.42	15.77
sd	0.40	0.50	0.46	0.46	3.26	0.30	0.43	0.40	0.25	1.06	3.85
median	2	2	2	2	8	2	2	2	2	8	16
IQR	0	1	1	1	2	0	0	0	0	1	2
range	1-2	1-2	1-2	1-2	4-15	1-2	1-2	1-2	1-2	4-8	8-23

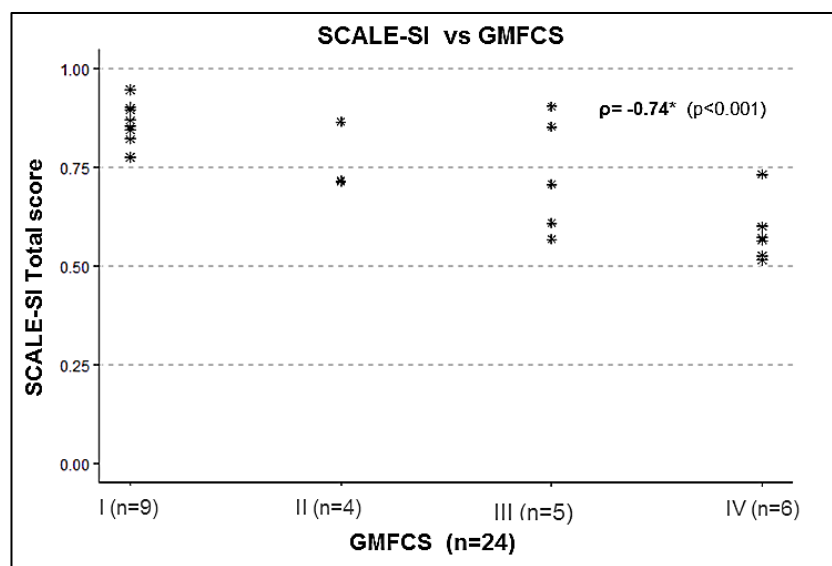
Abbreviations: SCALE-SI: Selective Control Assessment of the Lower Extremity – Similarity Index; sd: standard deviation; IQR: Interquartile Range.

6.4.1 Validity

Spearman-rank correlations between the SCALE-SI and the SCALE joint pairs were moderate to strong, with the exception of two weak and non-significant correlation pairs (for the less affected ankle and the more affected knee joint). The correlations for the summed leg scores were strong (less affected leg: $\rho = 0.75$; more affected leg: $\rho = 0.78$; for both: $p < 0.001$) (Table 13). A high correlation was found for the total scores ($\rho = 0.90$, $p < 0.001$) (Table 13). There was a high (negative) correlation between the total SCALE-SI and the GMFCS ($\rho = -0.74$, $p < 0.001$) (Figure 13).

Table 13: Study 4 - Correlations SCALE scores vs SCALE SI

Correlations SCALE-scores vs SCALE-SI	less affected side		more affected side	
	Spearman-rank correlation	p-value	Spearman-rank correlation	p-value
Hip	$\rho=0.53^*$	0.00825	$\rho=0.45^*$	0.02786
Knee	$\rho=0.61^*$	0.00161	$\rho=0.35$	0.09291
Ankle	$\rho=0.33$	0.1093	$\rho=0.56^*$	0.00394
STJ	$\rho=0.67^*$	0.00033	$\rho=0.55^*$	0.00513
Leg	$\rho=0.75^*$	<0.001	$\rho=0.78^*$	<0.001
Total score	$\rho=0.9^*$ ($p<0.001$)			

**Figure 12:** Study 4 - Correlation total SCALE score and GMFCS levels

Abbreviations: SCALE-SI: Selective Control Assessment of the Lower Extremity – Similarity Index; GMFCS: Gross Motor Function Classification System

Concerning discriminative validity, all SCALE-SI scores (separate joints, legs and total) were significant lower (Mann Whitney U test total score: $U = 87.000$, $p < 0.001$, $r = .65$) in children with CP when compared to their neurological intact peers (Table 14). There were no significant differences in the SCALE-SI between the four joint-pairs, neither for the less nor for the more affected leg (Friedman test, Chi-square value of 4.950, $p = 0.18$ for the less affected side; Chi-square value of 3.00 $p = 0.39$) (Fig 13). There was no statistically significant difference between the more and less affected leg in children with bilateral diagnosis (Wilcoxon Signed Rank Test: $Z = -0.971$, $p = 0.64$, $r = .217$), but one for children with unilateral involvement (Wilcoxon Signed Rank Test: $Z = -2.214$, $p = 0.027$, $r = -.904$) (Fig 14).

Table 14: Study 4 - Differences neurologically intact children vs. children with CP

Mann Withney U Test SCALE-SI		less affected side		more affected side	
		NI	CP	NI	CP
Hip	median (IQR)	0.98 (0.10)	0.76 (0.30)	0.97 (0.11)	0.77 (0.19)
	p-value	p<0.001*		p<0.001*	
Knee	median (IQR)	0.95 (0.12)	0.78 (0.29)	0.94 (0.10)	0.74 (0.36)
	p-value	p=0.001*		p=0.005*	
Ankle	median (IQR)	0.97 (0.06)	0.93 (0.20)	0.97 (0.11)	0.80 (0.39)
	p-value	p=0.04*		p<0.001*	
STJ	median (IQR)	0.94 (0.05)	0.89 (0.26)	0.95 (0.05)	0.63 (0.53)
	p-value	p=0.02*		p<0.001*	
Leg	median (IQR)	0.94 (0.00)	0.83 (0.17)	0.93 (0.00)	0.71 (0.19)
	p-value	p<0.001*		p<0.001*	
Total score	median (IQR)	NI = 0.93 (0.00)		CP = 0.75 (0.26)	
	p-value	p<0.001*			

Abbreviations: SCALE-SI: Selective Control Assessment of the Lower Extremity – Similarity Index, STJ: SubTalar Joint; NI: Neurologically Intact children; CP: Cerebral Palsy; sd: standard deviation

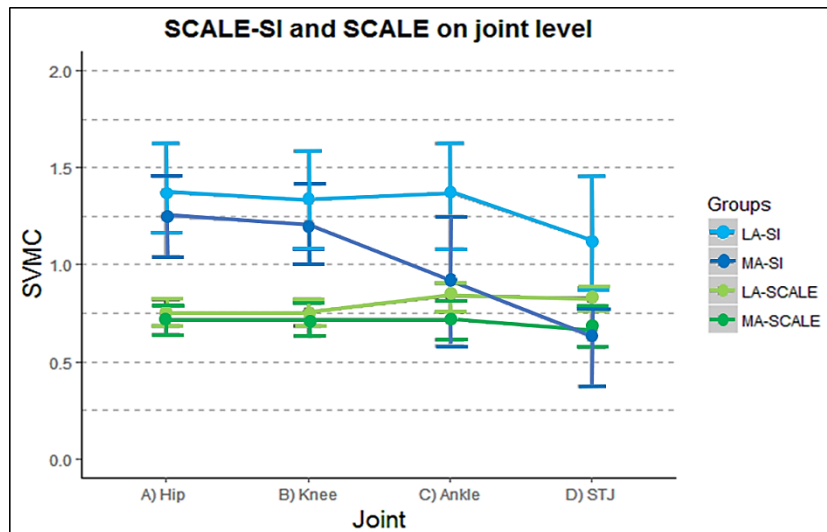


Figure 13: Study 4 – Hypotheses proximal-distal-concordance:

the tested joints (hip, knee, ankle, STJ) are displayed on the x-axes while the participant's level of SVM (mean SCALE joint score in blue and mean SCALE-SI in green) is displayed on the y-axes.

Abbreviations: SCALE-SI: Selective Control Assessment of the Lower Extremity – Similarity Index, STJ: SubTalar Joint; LA-SI: SCALE-SI for the less affected limb; MA-SI: SCALE-SI for the more affected limb; LA-SCALE: SCALE of the less affected limb; MA-SCALE: SCALE of the more affected limb

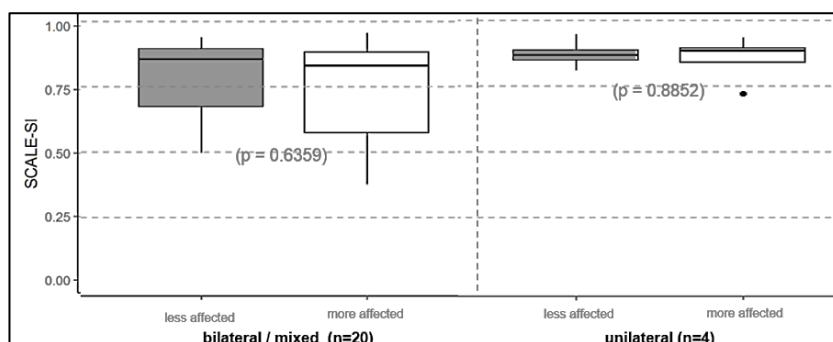


Figure 14: Study 4 – SCALE-SI grouped by leg involvement and diagnosis:

bilateral involvement n=20; unilateral involvement n=4; grey shadow box plot = less affected leg, white box plot = more affected leg.

Abbreviations: SCALE-SI: Selective Control Assessment of the Lower Extremity – Similarity Index

6.4.2 Reliability

With ICC values exceeding 0.9 for the total, 0.8 for the limb and 0.7 for the joint SCALE-SI scores and the lower limits of the 95% CI all above 0.4, the test-retest reliability was moderate to good. The MDC varied between 0.14 (total score) and 0.42 (more affected ankle) (Table 15). Bland Altman plots for SCALE-SI joint and leg levels are shown in Figure 15 for the less affected side and in Figure 16 for the more affected side.

Table 15: Study 4 - Relative and absolute reliability SCALE-SI

		Children with spastic CP (n=24)									
		less affected leg					more affected leg				
		hip	knee	ankle	STJ	leg	hip	knee	ankle	STJ	leg
Descriptive	Mean (SD)1	0.75 (0.18)	0.75 (0.19)	0.82 (0.20)	0.81 (0.16)	0.78 (0.14)	0.69 (0.19)	0.70 (0.19)	0.69 (0.26)	0.63 (0.23)	0.68 (0.16)
	Mean (SD)2	0.70 (0.22)	0.77 (0.16)	0.83 (0.19)	0.80 (0.17)	0.78 (0.15)	0.68 (0.18)	0.72 (0.20)	0.67 (0.27)	0.63 (0.24)	0.68 (0.18)
Relative Reliability	ICC (2.1)	0.77	0.77	0.72	0.70	0.86	0.80	0.76	0.74	0.84	0.88
	p value	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
	95% CI	0.51 - 0.90	0.51 - 0.90	0.42 - 0.87	0.40 - 0.87	0.66 - 0.94	0.57 - 0.91	0.50 - 0.90	0.47 - 0.89	0.64 - 0.93	0.72 - 0.95
Absolute Reliability	SEM	0.12	0.12	0.15	0.14	0.09	0.12	0.15	0.14	0.10	0.08
	MDC	0.35	0.33	0.40	0.38	0.26	0.32	0.42	0.40	0.26	0.23

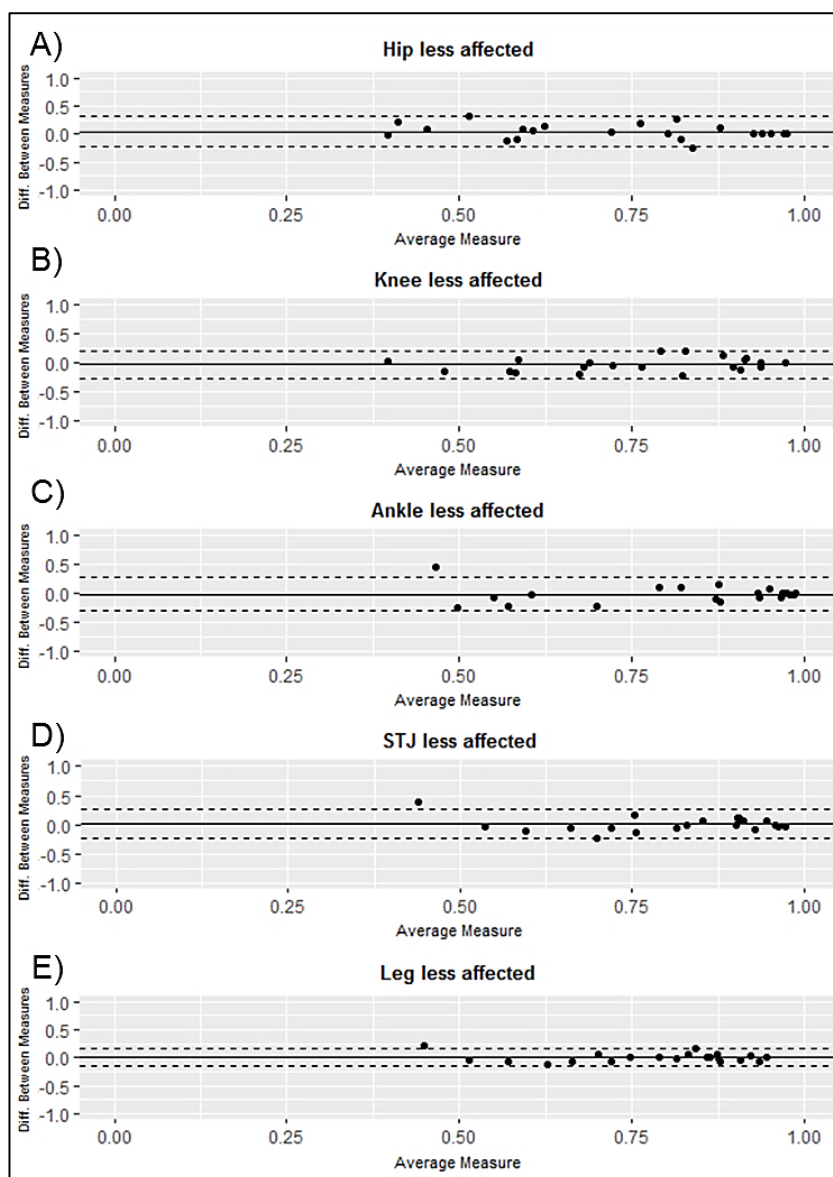


Figure 15: Study 4 - Bland-Altman Plots: Test retest reliability SCALE-SI – less affected side

● less affected joint children with CP (n=24), Rows: A) Hip joint; B) Knee joint; C) Ankle joint; D) STJ joint; E) leg

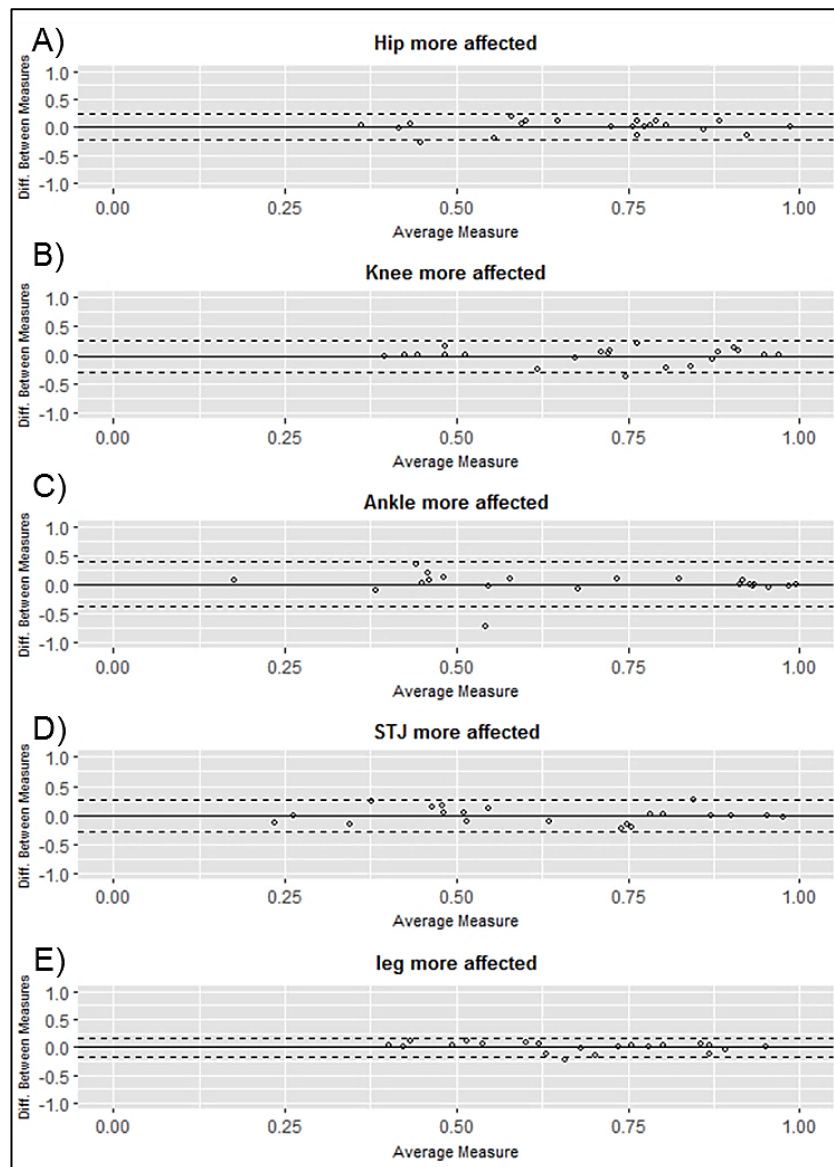


Figure 16: Study 4 - Bland-Altman Plots: Test retest reliability SCALE-SI - more affected side:

○ more affected joint children with CP (n=24), Rows: A) Hip joint; B) Knee joint; C) Ankle joint; D) STJ joint; E) leg

Abbreviations: ICC, Intra-class correlation coefficient; CI, Confidence Interval; SD, Standard Deviation; SEM, Standard Error of Measurement; MDC, minimum detectable change. Please note, we could include data from only 21 participants in the reliability analysis.

6.5 Discussion

The results demonstrate that SVMC in children with CP can be quantified as the degree of similarity between the sEMG pattern of ten muscles in a person with CP and that of a group of neurologically intact adults. Thereby showing that combining the SCALE testing procedures with sEMG is an objective and probably more sensitive approach than the clinical assessment without sEMG recordings.

6.5.1 Validity

The SCALE-SI concurrent validity as assessed by its correlation with the SCALE was good for the sum scores (less affected leg: $p = 0.75$; more affected leg: $p = 0.78$; for both: $p < 0.001$), but less for the individual joint scores. The weaker correlation at the joint level might be explained by the fact that there were some children with CP whose SVMC joint performance was scored as “unable/0” by the SCALE, but their sEMG muscle-activation patterns were adequate although within a very low amplitude. As there were also children with low SCALE scores and less physiological sEMG muscle-activation patterns, these findings indicate the ability of the SCALE-SI to detect SVMC sensitively and independently from muscle strength. Exactly this increased sensitivity in measuring SVMC, especially within the lower SCALE levels (0 and 1), was the aim when introducing the SCALE-SI as a new measure.

However, the strong association of the SCALE-SI sum scores with the SCALE sum-scores (e.g. sum scores per leg or the total score) and with the GMFCS confirms the validity of this new measure. The high negative correlation with the GMFCS is also in line with previous results of the original SCALE and confirms our concurrent validity hypothesis that children with higher SVMC have better mobility (Kusumoto et al. 2016; Balzer et al. 2016; Fowler et al. 2009).

Our first discriminative validity hypothesis was that the SCALE-SI scores of children with CP would be statistically significant from those of neurological intact children. This hypothesis was confirmed. To the author's knowledge, this study was the first, which applied the SI to a neurological impaired population other than those studies that applied the SI after a stroke or spinal cord injury. By acquiring comparable reference-type muscle activation patterns from neurologically intact adults as well as from peers, SCALE-SI discriminative validity was established. As SVMC is known to be maturation-dependent, data from neurological intact adults were collected to gain norm values from individuals, in whom SVMC is known to be fully and optimally developed. In contrast, the comparison with neurological intact peers aimed to reflect on the developmental component when measuring SVMC. Our results confirm the developmental aspect of SVMC (Cahill-Rowley and Rose 2014; Staudt et al. 2003; Rothwell 1987). While in the adult reference population, all SCALE-SI scores fall between 1 and 0.75, this range widens to 0.5 in neurological intact peers until the age of 10 years (Figure 17).

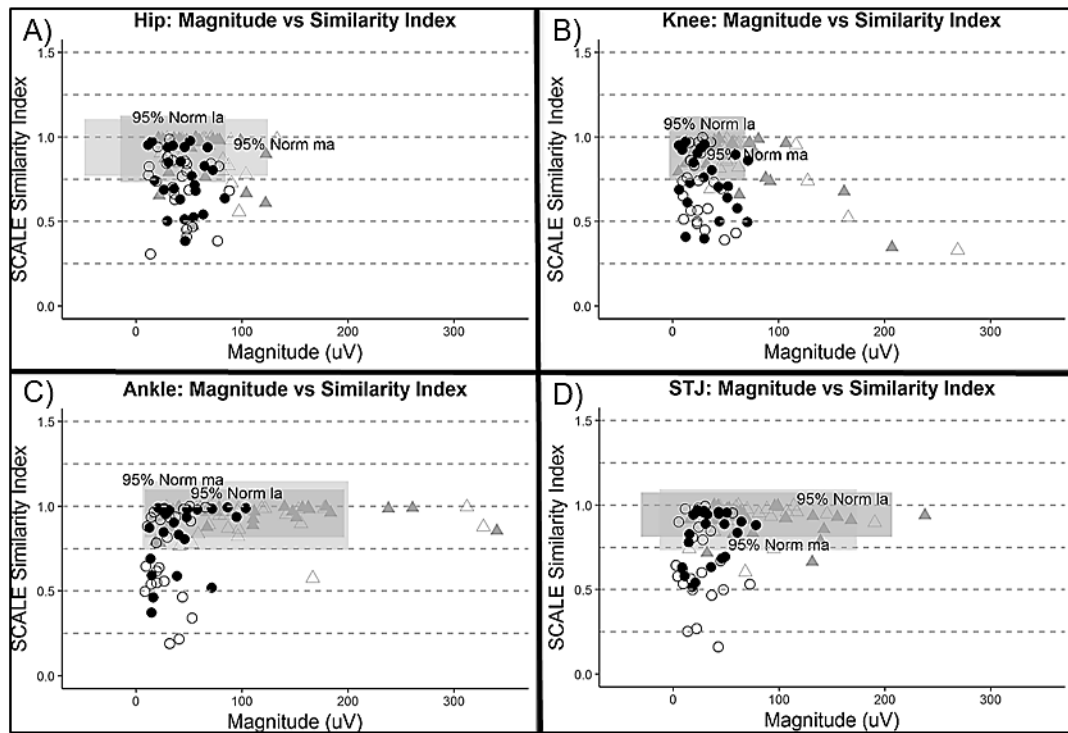


Figure 17: Study 4 - SCALE-SI:

Magnitude vs. Similarity plots for the A) hip, B) knee, C) ankle and D) STJ joint. The magnitude (the length of the multichannel EMG based-vector) is presented on the x-axis, while the SI index (the direction of the multichannel EMG based-vector) is displayed on the y-axis. Please note that the magnitude is a non-normalized value and displays the level of activation of the muscles, while the SI values display the similarity of the activation pattern in comparison to the norm. This figure shows that children with CP (circular shape) differ in both values in comparison to their peers (triangular shape).

- less affected joint in children with CP (n=24); ○ more affected joint in children with CP (n=24);
- ▲ less affected = dominant joint in neurologically intact children (n=32); △ more affected = non dominant joint in neurologically intact children (n=32);
- 95% Confidence Interval Norm (neurologically intact adults), less affected (la) = dominant joint (n=31);
- ▒ 95% Confidence Interval Norm (neurologically intact adults) more affected (ma) = non dominant joint (n=31)

Abbreviations: SI = Similarity Index; STJ = SubTalar Joint

Only one control child (12,11 years, month) showed voluntary contralateral gastrocnemius activation during knee joint movement, which was undetected by the tester during sEMG recording, but detected during offline sEMG-data processing. Age-dependency was also detected when scoring these children with the original SCALE (Figure 18). Above the age of ten years, almost all neurological intact children fall within the 95%CI of the adult population (Figure 19). Interestingly, the range of the sEMG magnitude, which is the second value of the VRI, was lower in children with CP (maximum of 100 μ V) compared to neurological intact participants (maximum of 300 μ V). This finding might be related to the smaller muscle size and volume in children with CP (Shortland 2011).

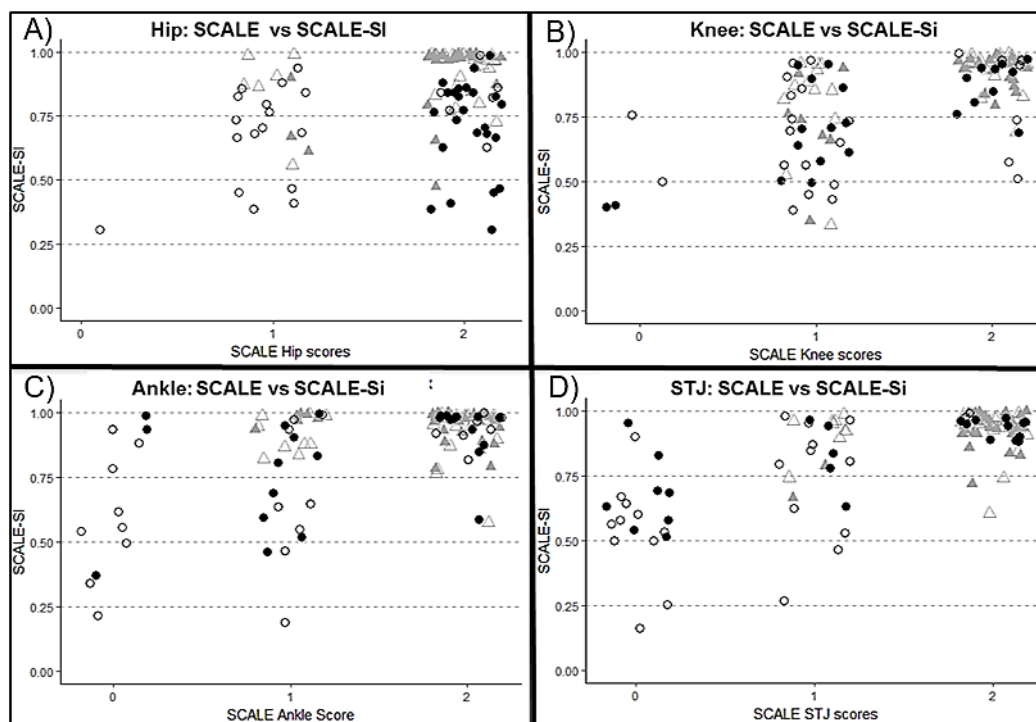


Figure 18: Study 4 - SCALE vs. SCALE-SI

for the A) hip, B) knee, C) ankle and D) STJ joint in children with CP and neurologically intact children. SCALE joint scores are plotted on the x-axis, while the SCALE-SI is displayed on the y-axis. This graph shows that the majority of neurologically intact children (triangular shape) had a SCALE score of 2 in combination with high SCALE-SI scores.

● less less affected joint in children with CP (n=24); ○ more affected joint in children with CP (n=24),
▲ less affected joint = dominant in neurologically intact children (n=32); △ more affected joint = non-dominant in neurologically intact children (n=32)

Abbreviations: SI = Similarity Index; STJ = SubTalar Joint

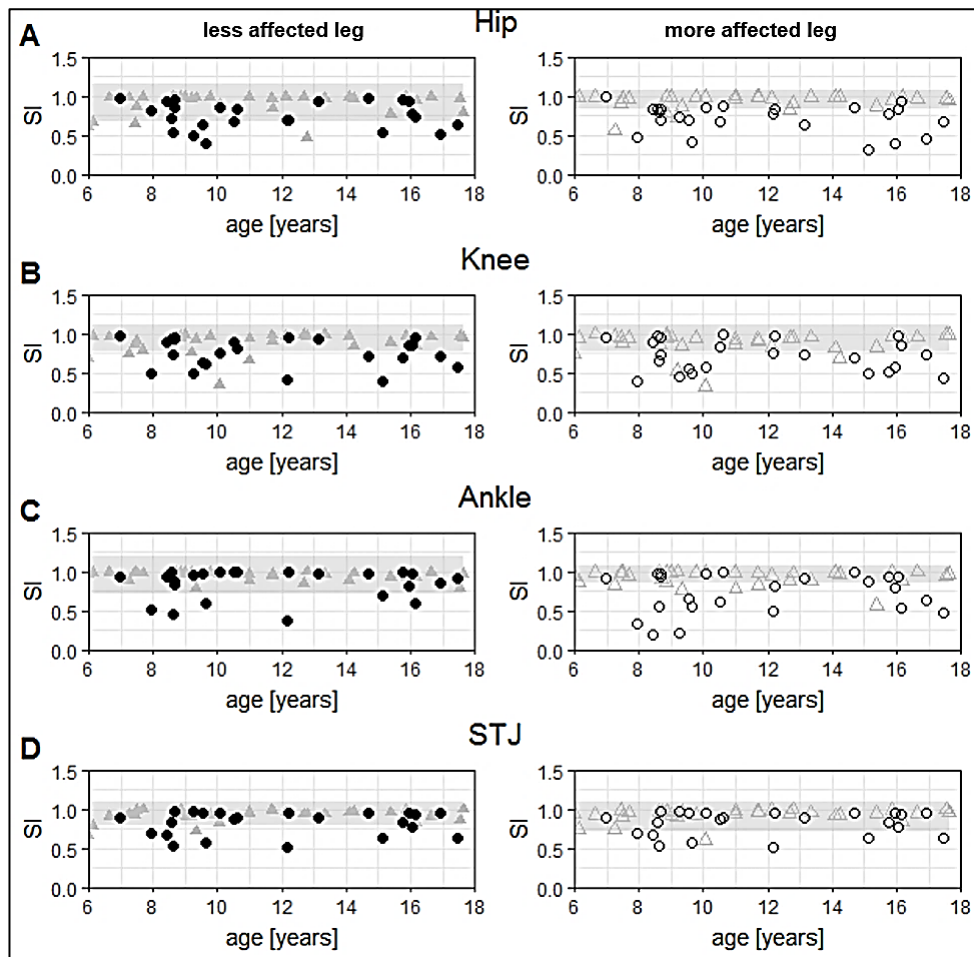


Figure 19: SCALE-SI vs. age

in children with cerebral palsy (circular shape) and neurologically intact children (triangular shape) for the less affected leg (first column) and the more affected leg (second column), for the A) hip, B) knee, C) ankle, D) STJ joint. Age is displayed on the x-axis and SCALE-SI on the y-axis, showing age dependent development of SVMC. The light grey bar represents the norm SCALE-SI values around that age, showing that after the age of 10 years the majority of neurologically intact children (triangular shape) falls within this normal range.

● less affected joint in children with CP (n=24); ○ more affected joint in children with CP (n=24); ▲ less affected = dominant joint in neurologically intact children (n=32); △ more affected = non-dominant joint in neurologically intact children (n=32); ■ 95% Confidence Interval Norm (neurologically intact adults)

Abbreviations: SI = Similarity Index; STJ = SubTalar Joint

Our second discriminative validity hypothesis, regarding the proximal-distal concordance of the SCALE-SI, could not be confirmed. These findings are in contrast to previous papers which have shown this proximal-distal concordance of the original SCALE (Balzer et al. 2016; Fowler et al. 2010). There was only a trend towards a more distal impairment in the more affected leg when quantified with the SCALE-SI. However, unlike the results of the previous studies, we noticed the same in our sample when evaluating the clinical SCALE scores (Figure 13). Especially the hip SCALE-SI score did not differ from the adjacent knee joint. This might be related to the altered

testing-position in sitting position (to minimize positional changes in relation to the sEMG). Standardization of the testing movement was difficult, as this position allowed movements around all three lower extremity joints. Furthermore, the SCALE-SI's diminished trend in proximal to distal concordance might be also explained by the fact that the SCALE-SI is measuring a slightly different constructs of SVMC in comparison to the SCALE. On the one hand, the SCALE-SI measures SVMC due to its sEMG sensitivity, movement (muscle strength) independently. On the other hand, due to the applied SI-algorithm, the SCALE-SI incorporates selective as well as pathological muscle activation within its equation, while the SCALE only relies on observational assessment of the selective movement. Both of these two specific characteristics of SCALE-SI are discussed within the limitations and clinical implications in more detail.

In relation to the lack of significant differences between the less and more affected limb for participants with unilateral involvement ($n=4$ in our sample), future studies should investigate these differences with much more participants.

In addition to the correlation between the SCALE-SI and the GMFCS, differences of SCALE-SI across the GMFCS levels and between adjacent levels were computed. The Kruskal Wallis test revealed a significant difference when comparing SCALE-SI values between participants with different GMFCS levels. However, probably due to the limited sample size and the additional post-hoc statistical corrections for multiple comparisons, statistically significant differences were only found between GMFCS levels I and II.

Future studies with a larger sample size should calculate clinical meaningful SCALE-SI threshold values for the GMFCS, by the way of a ROC curve analysis, comparable to the one from Lim and colleagues (2005) in patients with spinal cord injury and their American Spinal Injury Association level.

6.5.2 Reliability

An ICC value of 0.9 indicates that the total score of the SCALE-SI is highly reproducible and can therefore be recommended for clinical and scientific use. It is currently not possible to evaluate whether a MDC of 0.26 and 0.23 (less and more affected leg, respectively) is clinically acceptable, as realistic change estimates for SCALE-SI of the lower extremity are currently lacking.

6.5.3 Limitations and methodological considerations

The following methodological limitations should be taken into account when interpreting the results of this study.

The first three limitations consider the main limitations in measuring SVMC using the SCALE-SI. The first two limitations of the SCALE-SI relate to adaptations of the test procedure aimed at minimizing movement artefacts and acquiring appropriate muscle activity cut-off points. The last limitation regards the underlying SI algorithm of the SCALE-SI. Firstly, we changed the original SCALE testing position of the hip from side lying into sitting. This was done to minimize movement artefacts in the sEMG signals. Nevertheless, as in this testing position the leg is moved against gravity, a higher amount of muscle strength is needed, and the test condition is more difficult to standardise. Accordingly, this altered hip position might be not optimal for testing SVMC.

Secondly, to correct background levels of muscle activation during testing, we had originally planned to subtract the level of background sEMG activity measured during the initial baseline period from sEMG levels measured during the test conditions. In neurologically intact participants as well as in children with CP, the background muscle activity during baseline was in some muscles higher than during the test-trial, thereby generating negative values. This phenomenon has been observed previously (Fowler 2010; Lim et al. 2005) and reflects the difficulty in voluntarily relaxing the muscles. Hyperreflexia and/or spasticity might have increased baseline muscle activity additionally in the patients with CP. Although this procedure might be appropriate in neurologically intact adults, its appropriateness when measuring patients and especially children with positive UMN signs should be further established. Therefore, we decided to subtract the lowest amount of sEMG activity that was observed during the test-trial.

Thirdly, due to the SI algorithm, the SCALE-SI does not distinguish between difficulties in desired selective muscle activation and pathological co-activation of ipsi- or contralateral muscles. As such, SCALE-SI is an unweighted index, as activation of each muscle “weights” (counts) equally. Furthermore, the number of muscles resulting in possible co-activation is much higher than the primary activated ones. In order to reduce co-activation, the testing positions were selected and adapted (i.e. leg pedestal) in such a way that it would, normally, hardly require any activation of

uninvolved muscles. Nevertheless, the SCALE-SI is unable to differentiate between mirror movements and co-movements. The SI algorithm can only quantify the similarity of the muscle activation pattern of a participant to that of a 'normal population'.

Furthermore, the finding that the proximal-distal concordance is less prominent in the results of the SCALE-SI compared to the SCALE may also be explained by the fact that the SCALE-SI in comparison to the SCALE measures a slightly different construct of SVMC. Firstly, the SCALE-SI measures SVMC independent from observable joint movement (and thus muscle strength) on which the SCALE is based. Secondly, the SCALE-SI incorporates selective as well as pathological muscle activation within its equation, while the SCALE only relies on observational assessment of the selective movement. Both of these two specific characteristics of the SCALE-SI are discussed in more detail in section 6.2.

Lastly, although we aimed for a sample size of at least 30 participants, data of only 24 participants could be collected. The lack of recruitment was mainly caused by the following two issues. Firstly, pilot testing was prolonged to best possibly optimize the measurement procedure. Secondly, as the author's research group developed a pendant SVMC measure for the upper extremity (Selective Control of the Upper Extremity Scale- (SCUES-SI)), recruiting 30 participants, who were willing to be assessed twice, for both studies, was challenging within the given time frame.

6.5.4 Clinical Implications and future work

One major advantage of the SCALE-SI above the SCALE is its joint-movement and muscle strength independency, therefore allowing to measure SVMC also in very weak participants. While the SCALE rates absence of joint movement in the tested joint as 'unable SVMC (0)', the SCALE-SI could distinguish between participants with no movement but relatively normal physiological muscle activation patterns and those with no movement but also less physiologically and thereby more impaired SVMC. The sensitivity of the SCALE-SI in detecting physiological and impaired SVMC in participants with and without joint-movement might indicate its future application as a prognostic tool to optimize physiotherapy training plans. For instance, while training SVMC (e.g. by isolated joint movement training or functional training (e.g. FES)) might be indicated in children with CP with a certain level of SVMC (e.g. SCALE-SI above

0.75); a more functional training, solely focusing on learning or persevering an activity (e.g. walking) might be the therapy of choice, when the SCALE-SI is low.

Besides the SCALE-SI's first positive validity and reliability findings, assessment of its responsiveness is required before it can be recommended as a SVMC outcome measure. Also, clinically important changes in relation to daily activities should be established to further investigate the clinical relevance of the SCALE-SI measure for the patients.

6.6 Conclusion

In conclusion, these first psychometric testing results of the SCALE-SI serve as reference values when applying the SCALE-SI in future clinical and scientific practice. It shows that carrying out the SCALE procedure with simultaneously multichannel sEMG application of the lower extremity in children with CP is possible with a few adaptations of the original SCALE testing protocol. The three main characteristics of the SCALE-SI can be summarized as followed: i) the SCALE-SI measures SVMC independently from joint movement; ii) the algorithm of the SCALE-SI quantifies SVMC in comparison to the individual activation pattern of the norm, but cannot distinguish between selective, co- and mirror muscle activation; and iii) the SCALE-SI is a valid and reliable measure when using separate limb and total (both limbs) scores. To establish the application of the SCALE-SI as an outcome measure for detecting therapy-induced changes of SVMC in children with CP, its responsiveness needs to be evaluated first.

.

Chapter 7: Final Discussion

7.1 Purpose of chapter

This final chapter comprises a synopsis of all study findings and an integrated discussion of all results and study limitations within the overarching clinical and scientific neuro-paediatric context. First, main findings of each of the previous studies will be summarized to set a common basis for the following discussion (section 7.2). These findings will then be synthesized and discussed in their relation to the main purpose of this PhD project: Finding a sensitive and psychometrically robust evaluated SVMC outcome measure for the lower extremity in children with CP. Accordingly, the following subsections will discuss study findings and related considerations about validity (section 7.3.1), reliability (section 7.3.2), and clinical utility (section 7.3.3) of the SVMC outcome measures that were investigated in this thesis. Additionally, ideas for improving the psychometric properties of the measures in future studies will be presented. Section 7.4 presents some considerations for future responsiveness testing of the applied measures. Finally, methodological considerations, as well as further clinical implications concerning all studies are shown in section 7.5.

Please note, to maximize reader-friendliness of this chapter, the author decided to include new figures within this discussion. These figures synthesize results of different studies as well as additionally analysed data for the purpose of this overall discussion.

7.2 Summary of main findings

This chapter summarises the main findings of all four studies.

7.2.1 Main findings study 1

The aim of this study was to investigate the impact of SVMC, among other common lower extremity impairments and trunk control, on gait capacity in children with CP. Although SVMC was omitted from the final regression model for gait capacity, the results of this study showed that SVMC was strongly associated with gait capacity ($p = -0.68$), trunk control ($p = 0.76$), as well as muscle strength ($p = 0.85$). We assumed that the high correlations between SVMC, muscle strength and trunk control (multicollinearity) were the cause for the exclusion of SVMC from the final multiple

regression model. Of the three other independent variables, trunk control and spasticity each explained a different part of the variance of gait capacity and remained within the model ($R^2 = 0.67$). While a strong relationship between SVMC and gross motor function (van der Linden et al. 2018; Chruscikowski et al. 2017; Park and Kim 2013; Kim and Park 2011; Ross and Engsberg 2007) and muscle strength (Balzer et al. 2016) has also been established in previous studies, these were the first findings which revealed the association of SVMC to trunk control. The results from these studies show the interconnection between SVMC, muscle strength and trunk control for gait capacity in children with CP. Thereby they indicate the significance of these impairments in gait assessment and, potentially, rehabilitation.

7.2.2 Main findings study 2

The second study established validity and reliability properties of the German language SCALE version by confirming previous results (Fowler et al. 2009; 2010) as well as by providing new evidence. The SCALE's validity was confirmed by: i) a strong correlation between the SCALE and the FMA items ($p = 0.88$) ii) the GMFCS level ($p \leq -0.80$); as well as iii) by significant differences between the less and more affected limb in children with bi- and unilateral limb involvement and iv) between SCALE joint scores, showing a proximal-distal concordance. Additional hypotheses concerning the relationships of SVMC to muscle strength and spasticity were tested. While SVMC and muscle strength were strongly correlated ($p = 0.88$), its relationship with spasticity was lower ($p = -0.55$). This latter finding was discussed in relation to the incompletely understood pathophysiological nature of SVMC. Intra-class correlation coefficients (ICCs > 0.9) and MDC (below 2 points) for the inter- as well as the intra-rater reliability of the SCALE were within clinically acceptable ranges. These results support the regular clinical assessment of SVMC of the lower extremity in children with CP and might, in the long term, improve medical understanding and treatment of this central impairment. Nevertheless, due to both the SCALE's ordinal scaling and related lack of psychometric evidence on its responsiveness, its application is currently limited to diagnostic and prognostic purposes only.

7.2.3 Main findings study 3

This systematic review, which evaluated psychometric properties of SVMC outcome measures for the lower extremity in children with UMN lesions, identified the following four assessments: SCALE; SMC; m-Trost and Gillette's SMC test and three

laboratory-based measures, namely lower extremity joint kinematics, sEMG, and torque steadiness. Overall, the main findings revealed: i) a limited number of psychometric studies assessed the validity and reliability of SVMC outcome measures, especially for laboratory-based measures; ii) the absence of studies evaluating responsiveness for any measure, and iii) the SCALE as having the highest level of evidence regarding its reliability, validity, and clinical utility compared to other clinical and laboratory-based measures of SVMC. Nevertheless, the SCALE's application as a measure to assess therapy-induced changes in SVMC was questioned due to its broad ordinal scoring system. Accordingly, to augment the SCALE's ability to measure differences in SVMC in children with UMN lesions, it was suggested that it be used in combination with laboratory-based measures (e.g. sEMG or kinematics).

7.2.4 Main findings study 4

Consequently, study 4 dealt with validity- and reliability-testing of a sEMG based measure using the SCALE's measurement protocol. The SCALE-SI, which quantifies SVMC by comparing the similarity of a participant's multi-muscle sEMG pattern to that of an average control group (neurologically intact adults), was found to be an objective and more sensitive method compared to the ordinal rating system of the conventional SCALE. Its concurrent validity was evaluated by its correlation to the original SCALE scores as well as to the GMFCS levels. While these correlations were above 0.7 for summed leg and total scores for the SCALE and GMFCS, the correlations for the SCALE joint scores ($p=0.35 - 0.67$) were lower than expected. Although this finding, as well as the lack of the expected proximal-distal concordance of the SCALE-SI, were in contrast to the previously formulated hypotheses, they could be (partly) explained by the fact that SCALE and SCALE-SI are measuring slightly different constructs of SVMC. In contrast to the SCALE assessment, which relies on the observation of movement, the SCALE-SI measures SVMC by sEMG will detect muscle activity even without observable joint movement. Further, the SCALE-SI algorithm takes selective as well as pathological co-activation into account and is thereby not capable to differentiate between mirror movements and co-movements. The SCALE-SI only quantifies how the patient's multichannel sEMG activation pattern is – similar to normal. Concerning further discriminative results of the SCALE-SI, its validity was supported by the significant differences for all scores when comparing SVMC in children with CP and neurologically intact peers. Altogether, these results

show that combining the SCALE procedure with sEMG for quantifying SVMC of the lower extremity in children with CP is feasible and valid. When applying to test-retest situations, a change of the SCALE-SI total score above 0.14 can be considered as real change. However, to establish the SCALE-SI's application as an outcome measure for detecting therapy-induced changes of SVMC in children with CP, its responsiveness still needs to be evaluated.

7.3 Considerations when testing psychometric properties in SVMC measures

Only through outcome measures with robust psychometric properties, scientists and clinicians can accurately and precisely evaluate the efficacy of an intervention. The results of the systematic review of this PhD project have shown that such an 'ideal' tool is not yet available for measuring intervention effects on SVMC in children with CP. Therefore, this chapter will summarize evidence from the previous studies and suggest ideas for promoting the development of an 'ideal' SVMC outcome measure in the future. Accordingly, this section of the general discussion will synthesize the knowledge regarding psychometric properties of SVMC measures (chapter 2.3) with the previously described study results about their meaning for testing validity (section 7.3.1), reliability (section 7.3.2), and clinical utility (section 7.3.3) of SVMC measures. Each of these sections will start by briefly summarising the definition and main points of each psychometric property. This will be followed by discussion of the study results and experienced limitations in relation to this particular psychometric property. The last element of this section will present ideas and recommendations for studies aiming to investigate the responsiveness of the SCALE and SCALE-SI (section 7.3.4).

7.3.1 Considerations when testing validity of SVMC measure

In this section, the lessons learned from testing validity of the SCALE and the SCALE-SI will be described and discussed in greater detail. In accordance to the different subtypes of validity as defined by COSMIN (chapter 2.3.4), we consider in criterion validity (section 7.3.1.1) and construct validity (section 7.3.1.2) which includes both convergent and discriminant (divergent) validity. Within each of these subsections, the applied hypothesis, which was used to establish validity in relation to the SCALE and SCALE-SI, will be described and discussed. In section 7.3.1.1, the problem of identifying an adequate reference measure in relation to concurrent validity of the SCALE and SCALE-SI will be discussed. Concerning hypotheses testing, meaning

discriminant and convergent validity, the following topics will be discussed: establishing discriminative validity of SCALE-SI by i) comparing children with UMN lesions with neurological intact peers, ii) confirming proximal-distal concordance; and establishing convergent validity by using theoretical concepts about the relationship of SVMC and iii) spasticity, iv) muscle strength, v) trunk control, and vi) gross motor function (e.g. gait).

7.3.1.1 Considerations for criterion validity

Criterion validity is defined by the COSMIN panel as “the degree to which the score of a measurement instrument is an adequate reflection of a gold standard”. It is further subdivided into concurrent and predictive validity. Concurrent validity is demonstrated when a test correlates well with a measure that has previously been validated (but is not a gold standard), showing that both tests measure the same construct.

Ideally, the ‘gold standard’ is an objective measure which is 100% sensitive and specific concerning the presence and absence of a disease or disease state. In practice, for many measures, there are very few or no true ‘gold standard’ tests. Concurrent validity is tested by using an alternative pre-evaluated ‘gold standard’ test, measuring the similar or related construct of interest at the same time as the tool under validity evaluation (Mokkink et al. 2016). However, when comparing an instrument with a reference instrument that is not a ‘gold standard,’ it is unknown to which degree the correlation between instruments is influenced by the fact that both measures might not assess the exact same construct.

a) Concurrent validity of the SCALE

There is currently no gold standard for measuring SVMC. This might be related to the fact that SVMC can only be measured indirectly, as all related cortical activation patterns cannot be observed directly (section 2.3.6.1). Therefore, concurrent validity of the German version of the SCALE was established with the lower extremity items of the FMA (items III-IV) (study 2). The FMA is an internationally well-established assessment tool for stroke patients (Pandian et al. 2012). Item III tests if and how (0-2) the participant can move the knee and ankle in a position of combined synergies. Item IV requires more selectivity, by testing if and how the participant can selectively move the knee and ankle out of a synergy when standing with the hip at 0 degrees. Although these test items and the increased level of difficulty from item III to item IV

in the FMA are different when compared to the SCALE, the construct of both tests seems to be similar, namely the ability of the participant to voluntarily activate certain muscle(-groups) by simultaneously deactivating others (synergistic ones). This common convergent theoretical concept was supported by a statistically significant high correlation of $p=0.88$ between these two measures.

b) Concurrent validity of the SCALE-SI

To establish concurrent validity of the SCALE-SI, the previously validated SCALE was used. Although both tests rely on the same definition of SVMC and are assessed using a similar protocol, they measure slightly different constructs of SVMC. Firstly, while the SCALE is scored through the observation of movement and uses an ordinal scoring system, the SCALE-SI is derived from sEMG with a continuous value between 0 and 1. Secondly, due to the applied SI algorithm, the SCALE-SI is a measure which compares multi-muscle-activation pattern with the norm, but which cannot distinguish between the desired activation and pathological co-activation. In contrast, the SCALE focuses and rates how selective the joint movement was performed.

Because the SCALE-SI is unlike the SCALE derived from sEMG, it is able to measure the presence of SVMC even when no joint movement occurs. Although the original SCALE testing positions were developed to minimize the influence of muscle strength when measuring SVMC, its rating system is to some degree muscle strength dependent. As the SCALE assesses SVMC via joint movement, a minimum amount of muscle strength is required. For example, to assess the hip or ankle joint, a MMT of 2 (i.e. full range of motion without the influence of gravity) is needed or even a MMT of 3 (i.e. full range of motion against gravity), when testing, for example, the knee joint. The SCALE-SI, however, only requires a muscle strength of MMT 1 (i.e. muscle activation). This joint-movement and muscle strength independency of the SCALE-SI is considered as a major advantage of the SCALE-SI above the SCALE, as it allows for measuring SVMC in very weak participants. While the SCALE rates absence of joint movement in the tested joint as 'unable SVMC (0)', the SCALE-SI could distinguish between participants with no movement but relatively normal physiological muscle activation patterns and those with no movement but also less physiologically and thereby more impaired SVMC. Figure 13 shows a wide spread of SCALE-SI for SCALE 'unable/0' and 'impaired/1' scores. This increased sensitivity of the SCALE-SI was desired to be able to detect smaller differences of SVMC in children with CP. When SVMC was

scored as 'unable/0' by the SCALE, which was mostly the case around the ankle and the STJ joint, SCALE-SI was either low or high, depending on the selectivity of the unobservable muscle activation pattern of the participant. The sensitivity of the SCALE-SI in detecting physiological and impaired SVMC in participants with and without joint-movement might indicate its future application as prognostic tool to optimize physiotherapy training plans. For instance, while training SVMC (e.g. by isolated joint movement training or functional training (e.g. FES)) might be indicated in children with CP with a certain level of SVMC (e.g. SCALE-SI above 0.75), a more functional training, solely focusing on learning or persevering an activity (e.g. walking) might be the therapy of choice, when the SCALE-SI is low.

Considering the validity testing results of the SCALE-SI versus the SCALE, the increased sensitivity of the SCALE-SI may have resulted in the lower correlation on joint level. For instance, for the more affected knee and less affected ankle, correlations were lowest and not significant. Especially in these two joints, movement of four participants (in three of the knee and one for the ankle) looked physiological and were thereby rated as 'normal/2' by the SCALE, but the underlying muscle activation patterns were less physiological.

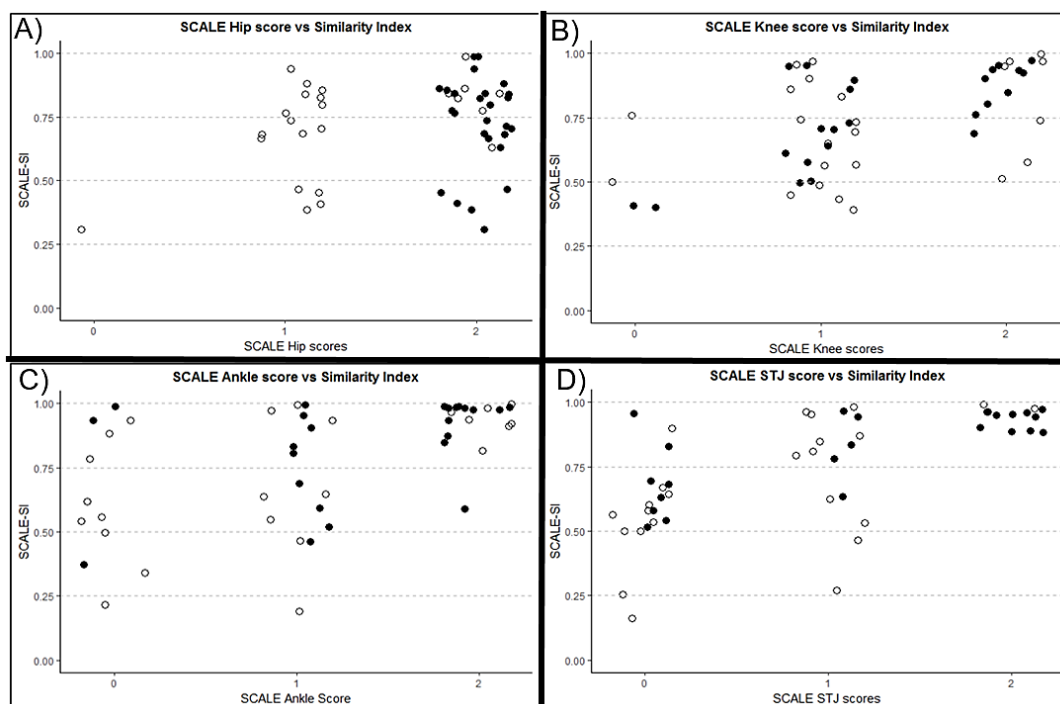


Figure 20: Concurrent validity: SCALE vs. SCALE-SI on joint level.

A) hip, B) knee, C) ankle and D) STJ in children with CP (circular shape). SVMC joint scores are displayed on the x-axis for the SCALE and for the SCALE-SI on the y-axis, showing the correlation between these two measures in terms of SCALE-SI's concurrent validity testing. Note: correlation for hip joint: less affected (la): $p=0.53^*$; more affected (ma) $p=0.45^*$; for knee joint: la: $p=0.61^*$; ma $p=0.35$; for ankle joint: la: $p=0.33$; ma $p=0.56^*$; for STJ joint: la: $p=0.67^*$; ma $p=0.55^*$; *: significant at $\alpha \leq 0.05$.

● less affected joint/leg children with CP (n=24) ○ more affected joint/leg children with CP (n=24);

Overall, these lower correlations at joint level between the two measures do not indicate that the SCALE-SI is less valid. It rather indicates that future SCALE-SI validity studies should select an alternative 'gold standard', which can measure SVMC more accurately than an ordinal-scaled assessment tool and which measures SVMC independently from joint movement, as well. Clinically, the ability of the SCALE-SI to measure SVMC independent from muscle strength (MMT1) also allows for measurement of changes in SVMC in more severely impaired children with CP (GMFCS III, IV) and it might allow identification for whom SVMC is worthwhile to train.

7.3.1.2 Considerations for hypotheses testing: convergent and discriminative validity

In cases without a 'gold standard', convergent and discriminative validity are possible alternatives. These two types of validity belong to the COSMIN category (box) 'hypotheses testing' (De Vet et al. 2011). When hypotheses testing is performed to evaluate validity (Appendix 1: Box F), *a priori* formulated hypotheses about the

relationship of the measurement scores and the scores of comparator instrument or group are tested. The expected relationship could either be that both tools measure the similar construct (*convergent validity*) or that the instruments are measuring different constructs (*discriminant and divergent validity*). (De Vet et al. 2011; Mokkink and Terwee 2010; Portney and Watkins 2000). Logically, the robustness of these two types of validity is strongly dependent on the previously defined construct of interest and the formulated hypotheses.

a) Suggestions for establishing discriminative validity by comparing SVMC in children with CP versus neurologically intact peers

For the evaluation of the discriminative validity of the SCALE-SI, we hypothesized in study 4 that children with CP would have lower SCALE-SI scores on joint, leg and total scores due to their corticospinal impairment. Thereby the aim was to explore the ability of the SCALE-SI in discriminating between these two groups, as well as between young (below ten years) and older (above ten years and under 18 years) neurologically intact children. We anticipated that younger children would also have lower SCALE-SI scores, as it has been shown that SVMC develops in line with the maturation process of the CNS until the age of 10 years (Cahill-Rowley and Rose 2014; Staudt et al. 2003; Rothwell 1987). Indeed, this age-dependency in neurologically intact children was shown by the data of both the SCALE-SI and even the ordinal SCALE. SCALE scores in the control group of young children were mainly scored as 'impaired', as they either showed co- or mirror movements (rated by the SCALE descriptors) (Appendix 13). For example, when testing the knee joint, co-movement of the foot was most frequently recorded and around the ankle or STJ joint; mirror movements of the contralateral foot frequently occurred. Considering the SCALE-SI scores of neurologically intact children, after the age of 10, scores for all joints fell within the 95%CI of the adult population (Appendix 12). To be able to compare SVMC of both groups of children with the most developed (i.e. adult) form of SVMC, the SCALE-SI prototype response vector was computed by the average muscle activity patterns of neurological intact adults. While in the adult reference population all SCALE-SI scores fall between 0.96 and 0.75, this range widens to 0.95 - 0.50 in neurologically intact children above the age of 10 years. The lower and upper borders of the 95% CI are even broader in children with CP where SCALE-SI scores for the less affected limb varied from 0.82 - 0.37 or for the more affected leg from 0.70 - 0.16. (Appendix 12). In relation to the age-dependency of SVMC, future studies with

an appropriate number of participants in all age groups, may apply an 'Analysis of Covariance' to control for age, when comparing SVMC in neurologically intact children and children with CP.

Validating the SCALE-SI via comparison between children with CP and neurologically intact peers has thereby supported the previously suggested age-dependency of SVMC. Furthermore, this kind of validation improved understanding of SVMC *per se* as reference values were missing up until now.

b) Suggestions for establishing discriminative validity by the concept of proximal-distal concordance (difference on joint level)

Another hypothesis, which was used to test discriminative validity of the SCALE (study 2) and the SCALE-SI (study 4), was that SVMC would be lower distally. This concept is also known as proximal-distal concordance and was first described by Brunnstrom in 1996 in patients with stroke and later shown by Fowler and colleagues (2010) in children with CP. Neurophysiologically, this phenomenon is often explained by the cortical representation of the lower extremity on the homunculus (Figure 3). While the motor cortex area for the hip and trunk lay at the medial midpoint of one hemisphere, the motor cortex area of the toes and ankle lies deepest within the periventricular area. As the periventricular area is most commonly injured in children with spastic CP, corticospinal fibers supplying the distal joints get more frequently injured than those supplying the knee and hip (Staudt 2007; Glenn et al. 2007).

While this proximal-distal concordance was observed when testing the German SCALE version (study 2), it was only marginally present for the SCALE scores in study 4, and even less for the SCALE-SI of the more affected leg (Figure 14). The following four issues might explain this difference in findings. Firstly, the assessment of SVMC at the hip joint was altered in study 4 because of the requirements of recording sEMG. As mentioned in paragraph 6.5.3, it was found that testing the hip joint in sitting position instead of in the originally described side-lying position was not optimal for evaluating SVMC, as the sitting position was more difficult to standardize. Furthermore, the required hip movement depended more on muscle strength, because the leg had to be flexed against gravity. This might explain why the SCALE hip joint scores in study 4 were lower than in study 2 and why differences between SCALE-SI hip and knee scores were smaller than it has been expected from the results of study 2.

Secondly, the differences in GMFCS levels in both studies might be another possible explanation why a clear proximal-distal concordance was only observed in study 2. While in study 2 children with GMFCS level I dominated (n=23/39), the severity levels in study 4 were almost equally distributed (GMFCS I and II, n: 13/24; GMFCS III and IV, n:11/24). As in more severely impaired children (GMFCS III and IV) SVMC tends to be impaired in all lower extremity joints, differences between the SCALE joints are absent. Therefore, the greater proportion of severely impaired children in study 4 might be another explanation why the proximal-distal-concordance was less obvious in study 4 compared to study 2.

Another possible explanation for the inconsistency with regard to proximal-distal-concordance of the two studies might be found in the fact that the SCALE-SI and SCALE are measuring slightly different aspects of SVMC. While the SCALE is dependent on a minimum of muscle strength (MMT 2-3) and an observable range of motion, the SCALE-SI measures SVMC independently from movement. As range of motion around distal joints is in general smaller compared to proximal ones (i.e. knee), differentiating between a SCALE score of 1 and 0 becomes progressively challenging. Especially if contractures are present, the movement range would become so small that it is almost impossible to observe whether the range of motion is above or below the active 50% range of motion, which would often result in a lower SCALE score at the more distal joints. Using sEMG for the SCALE-SI allows for the measurement of only slight but physiological muscle activation patterns around these distal joints. This increased sensitivity in detecting selective muscle activation of SCALE-SI in comparison to the SCALE especially around the ankle was also observed in relation to long-term orthosis treatment of many ambulatory children with CP.

Due to the stabilizing and limiting effect of the orthosis, the child's ability to learn to control and activate these distal muscles is very limited (learned-non-use). Afferent and proprioceptive input to the brain is decreased and might downregulate CTS connectivity and maturation especially during the early development stages of the central nervous system. As a functional consequence of this, muscles may become weak, although their physiological activation might be still possible. This will result in a lack of movement when assessed using the SCALE, but in a coherent sEMG activation pattern, as quantified by the SCALE-SI. Accordingly, as proximal-distal-concordance is more prominent in the SCALE assessment, this concept might be

more related to movement and muscle strength and not solely to SVMC and motor cortex activation.

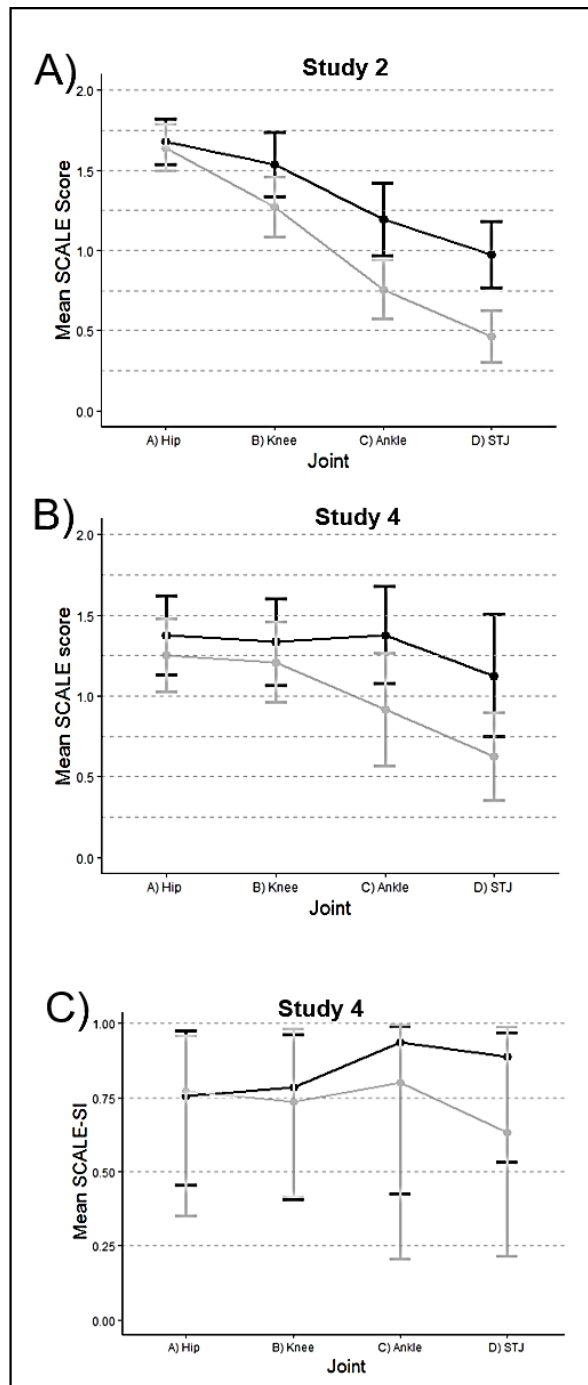


Figure 21: Proximal-distal concordance:

mean SCALE joint scores of A) study 2 n=39 and b) study 4 (n=24) and C) mean SCALE-SI joint scores of study 4 (n=24). Joints are displayed on the x-axis and their SVMC score (SCALE or SCALE-SI) on the y-axis for the less affected side in black and the more affected side in grey.

● / ---- less affected joint in children with CP; ● / ---- more affected joint in children with CP

Abbreviations: SCALE-SI: Selective Control Assessment of the Lower Extremity – Similarity Index; STJ: subtalar joint

c) Establishing discriminant validity: theoretical relationship between SVMC and spasticity

By testing hypotheses regarding the relationship between SVMC (SCALE and SCALE-SI) and spasticity, the aim was to establish discriminant convergent validity of the SCALE and the SCALE-SI. These a priori formulated hypotheses were based on the current pathophysiological understanding of SVMC. As stated in chapter 2.2.5, SVMC is one possible negative UMN sign in children with a lesion of the first motor neuron. In clinical practice, a mixture of another negative (e.g. muscle weakness) and additional positive (e.g. spasticity) UMN signs is often present in these patients (Carr 2014). Therefore, it is difficult to investigate positive and negative UMN signs separately. Their common pathogenesis and their interconnected hidden neural pathways complicate investigation of linear causality. Concerning SVMC and spasticity, there is an ongoing debate about the real pathophysiological nature of (decreased) SVMC (chapter 2.2.6). One opinion advocates a loss of selective control that is related to a loss of inhibition and the release of primitive flexor/extensor patterns and would thereby strongly relate to hypertonia (positive UMN sign). Another opinion argues that impaired SVMC is an impairment on its own (negative UMN sign), caused by the loss of connections to the CSTs. Finally, there is a third opinion that think of SVMC as the result of both negative and positive UMN signs.

Results of study 1, 2 and 4 of this PhD project seem to be in favour of the second and third theoretical construct regarding the nature of SVMC. Correlation results between the SCALE or the SCALE-SI and the MAS were only moderate negative ones ($p = -0.43$ (SCALE study 1) to -0.61 (SCALE-SI, study 4)), showing that there is a relationship, though not as strong as expected in relation to opinion one, in which impaired movement control is thought to appear purely due to impaired reflexes and hypertonia (Figure 15). Nevertheless, it should be noted that the overall level of spasticity was mild in all three samples, because there was a GMFCS level I to II dominance in study 1 and 2 and GMFCS level I-III dominance in study 4. Please note also that participants of study 1 ($n=52$) and study 2 ($n=39$) were partly the same ($n=30$). Thereby, a separate interpretation of the results of these two studies should be handled with great care. Nevertheless, their results were cautiously used to reinforce observed trends within the data of this PhD project.

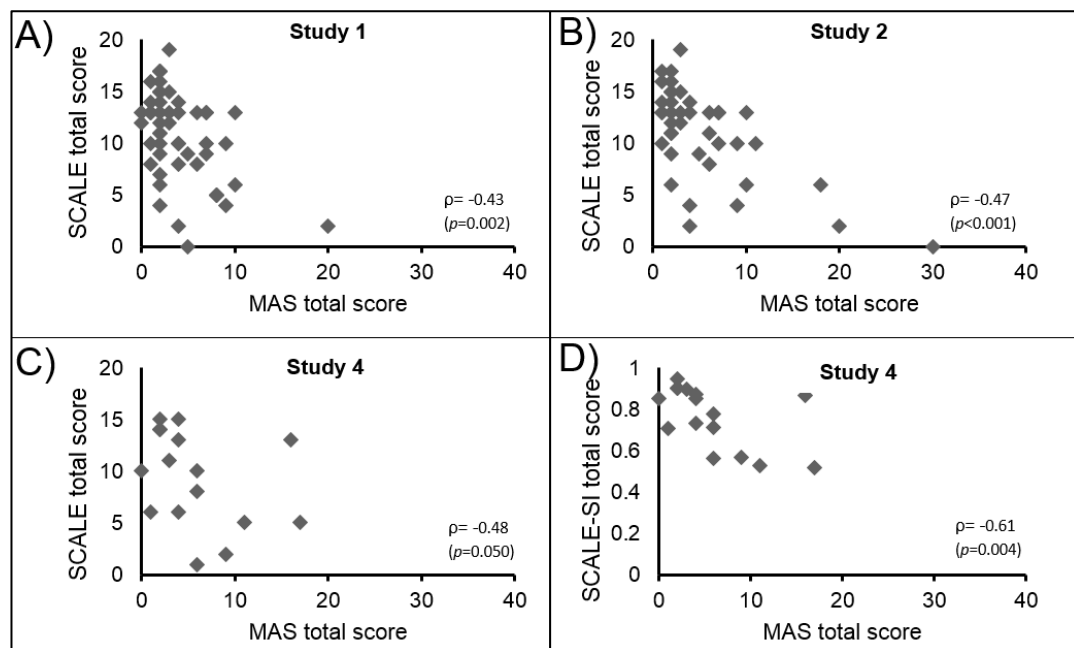


Figure 22: Correlations SCALE, SCALE-SI vs MAS:

A) correlations results SCALE vs MAS study 1 (n=52); B) correlations results SCALE vs MAS study 2 (n=39); C) correlations results SCALE vs MAS study 4 (n=24); D) correlations results SCALE-SI vs MAS study 4 (n=24)

Abbreviations: SCALE-SI: Selective Control Assessment of the Lower Extremity – Similarity Index; MAS: Modified Ashworth Scale

Results for the MAS show a large spread of SCALE and SCALE-SI scores in patients with low MAS. Nevertheless, this variety in SVMC in patients with low spasticity might indicate that in these participants, impaired SVMC is not a consequence of hypertonia, but should be considered as a negative UMN by itself. This impression changes when looking to the other end of the scaling system, where a high level of spasticity is associated with a low SVMC in children with spastic CP. Only one participant (study 4) was found, who had a relatively high level of spasticity and good SVMC (SCALE-SI). This participant (GMFCS IV) had contractures around the ankle and knee as well as a high level of spasticity. Thereby, the active range of motion was limited, but muscle activation was physiologically undisturbed, as seen by the sEMG. Therefore, when interpreting the correlation results of study 4, this outlier should be kept in mind, as correlations within small samples are known to be strongly affected by one or two extreme values.

Alternatively, these findings might also indicate that the MAS is too insensitive to assess hypertonia. This argument is supported by the current debate about whether there should be continued use of the MAS (Fleuren et al. 2010), due to its poor inter-rater reliability. Also, because its scoring can be influenced by several interfering

factors which are difficult to standardise, like the emotional state of the patient. Due to this questioned sensitivity of the MAS as well as due to the incompletely understood nature of SVMC and its interconnection with spasticity, the idea to test convergent validity of SVMC measures by its relation to spasticity, might not conducive to the establishment of this psychometric property. Future psychometric studies for SVMC measures might, therefore, consider using an alternative spasticity measurement tool with appropriate psychometric properties that measure spasticity more objectively (e.g. sEMG supported Pendulum Test (Fowler 2010) or an isokinetic dynamometry measure (Pierce et al. 2006)). The benefit of using an isokinetic strength measure is the addition of velocity as standardisable measurement component, allowing the testing of velocity-dependent increase in tonic stretch reflexes in a controlled manner. Previous studies, which used an isokinetic strength measure (Biodex) around the ankle and knee in children with CP, found that this measure is feasible and reliable, and capable of investigating the influence of spasticity on voluntary force production (Pierce et al. 2006; Damiano et al. 2001; Engsberg et al. 2000; Ayalon et al. 2000). Accordingly, an isokinetic strength measure might allow for testing the correlation between spasticity and SVMC as well as for muscle strength and SVMC.

In summary, the correlation results of these studies helped to gain a better understanding of the real nature of SVMC in children with CP. Larger cohort-studies are now needed, that include participants with a wider range of MAS values than reported in these studies, in order to facilitate a more conclusive understanding of the influence of spasticity on SVMC in children with CP.

d) Establishing convergent validity: theoretical relationship between SVMC and muscle strength

In contrast to spasticity, correlation coefficients between SCALE and muscle strength assessed by the MMT were strong ($p > 0.8$) in study 1 and 2. This could indicate that a correctly applied MMT will partially reflect the ability to activate a muscle (group) selectively. Unfortunately, the MMT was not used in study 4. This was done in order to keep measurement duration to a maximum of 1 hour to ensure compliance of the participants for validity and test-retest reliability testing. Although correlations for the MMT were strong and the MMT is commonly and easily applied in the clinical and scientific setting, it might not be the optimal reference tool for evaluating validity of a SVMC measure. Firstly, MMT's psychometric properties have not yet been

established for the lower extremities in children with CP. Secondly, its scaling system is quite broad in relation to its ordinal nature and does not assess the strength in different ranges of motion. An (instrumented) isokinetic strength measure might be a more appropriate comparison instrument, because it would allow the measurement of strength throughout the same range of movement, as with the measure of SVMC.. Additionally, it allows for the investigation of the influence of standardized movement speed on muscle strength and selective muscle activation. Establishing convergent validity of the SCALE-SI by an isokinetic-strength measure (Pierce et al. 2006; Damiano et al. 2001; Engsberg et al. 2000; Ayalon et al. 2000), as mentioned in the section above, might furthermore improve our understanding of the interrelated nature of these three UMN signs (impaired SVMC, muscle weakness and spasticity) in children with CP.

e) Establishing convergent validity: theoretical relationship between SVMC and trunk control

Referring back to Figure 3, showing the close neuroanatomical positions of the lower extremity and the trunk on the homunculus of the motor cortex led to the idea to investigate the relationship between these two motor control impairments. Although study 1 was not a psychometric study, interpretation of its correlation results between the SCALE and the TCMS might be extended to evaluate convergent validity of the SCALE within the context of this final discussion chapter.

As expected, correlations were strong between the trunk and the SCALE ($p=0.76$, $p<0.001$). This might indicate that the ability to control one's lower extremity and trunk are closely interrelated. A child with low trunk control will also have problems to control movement of the lower extremity. This interdependence is commonly observed within the clinical practice, but this is, to the author's knowledge, the first study systematically investigating this connection.

Future studies aiming to investigate convergent validity of the SCALE or SCALE-SI in relation to trunk control might select other references, i.e. more objective trunk-control measures, to further test the above stated hypotheses about SVMC of the lower limbs and the trunk. One idea would be the kinematic analysis of the centre of mass or pressure gained from a standardized trunk control force-plate measurement. Currently, our research group has developed such a force-plate based trunk control measurement tool, which tests trunk control playfully by means of a computer game.

The game is played with the participants sitting upright and without them having foot contact on the force plate. The child is asked to selectively follow a predefined path by moving the trunk in different directions as well as to catch and hold specific target points by holding the trunk posture, respectively. Theoretically, within the context of the SCALE-SI's validity testing, one would expect that children who are able to follow the predefined path more accurately and who are able at catching and holding specific target points, also have better motor control in their lower extremity.

f) Establishing convergent validity: theoretical relationship between SVMC and gross motor function skills

Finally, the possibility to establish convergent as well as discriminative validity of the SCALE or SCALE-SI via its relation to a gross motor skill measures will be discussed. The clinical observation behind this concept is that children with a higher level of mobility tend to have better movement control. This hypothesis was confirmed by study 1 via a moderate negative correlation between the SCALE and mTUG ($\rho = -0.69$; $p < 0.001$), as well as by study 2 and 4 via a strong negative correlation between the SCALE respectively SCALE-SI and the GMFCS ($\rho = -0.83$ and $\rho = -0.74$; $p < 0.001$). In relation to this discussion, the correlations between SCALE and GMFCS were also computed for study 1 and study 4. All correlations were equal or higher than $\rho = -0.70$ (Figure 16). These high correlations between SVMC and gross motor functioning are furthermore confirmed by other studies (Chruscikowski et al. 2017; Y. Kusumoto et al. 2016; Fowler et al. 2009) and show the importance of SVMC on the ICF activity level.

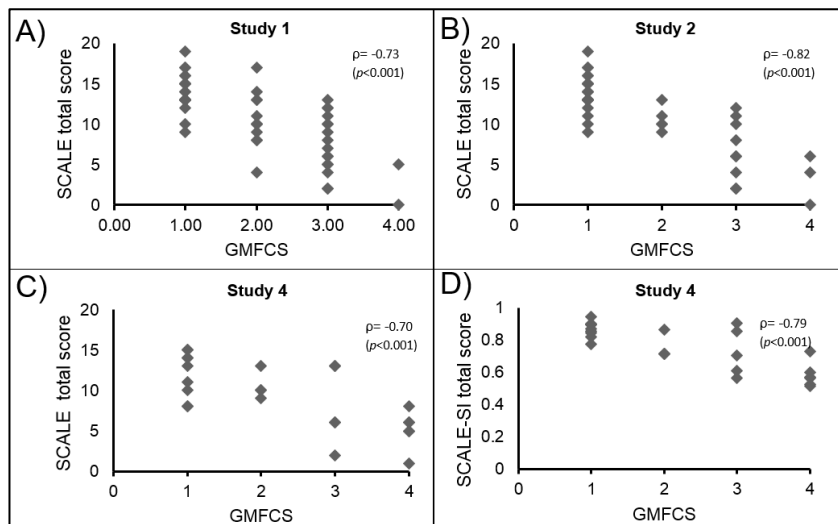


Figure 23: Correlations between SCALE / SCALE-SI and GMFCS:

A) correlations results SCALE vs GMFCS study 1 (n=52); B) correlations results SCALE vs GMFCS study 2 (n=39); C) correlations results SCALE vs GMFCS study 4 (n=24); D) correlations results SCALE-SI vs GMFCS study 4 (n=24)

Abbreviations: SCALE-SI: Selective Control Assessment of the Lower Extremity – Similarity Index; GMFCS: Gross Motor Functioning Classification

In addition to the correlation between the SCALE/SCALE-SI and the GMFCS, differences of SCALE/SCALE-SI across all GMFCS levels and between adjacent levels were computed in study 2 and 4 to establish discriminative validity. Overall, in both studies, the Kruskal Wallis Tests revealed only a significant difference between GMFCS level I and II. This was probably due to the limited sample size as well as to the unequal distribution within each GMFCS level. Performing a power analysis (80% power, two-tailed alpha .05) revealed that a sample size of at least 19 participants in GMFCS level II and III and 29 participants in GMFCS level III and IV would be required to determine statistically significant differences between these GMFCS levels.

Furthermore, future studies might be advised to employ a different mobility measure for establishing validity of the SCALE. A recent study showed that there is a moderate correlation ($\rho = -0.603$) between the Gait Profile Score, which is an index of how different a person's gait is compared to neurologically intact persons and the SCALE-score (Chruscikowski et al. 2017). In line with the results of this cohort-study, previous studies already demonstrated the relationship between the SCALE and specific gait parameters. Another interesting finding was shown by a recent study about the influence of SVMC among other lower limb impairments on 'RaceRunning' performances in athletes with hypertonia, ataxia or athetosis. The correlation between 100m 'RaceRunning' speed and SCALE ($\rho = 0.47$) was lower than when compared

to gait, but it was still statistically significant. Interestingly, the strongest association between the 'RaceRunning' speed and the other lower extremity impairment was found with spasticity. This higher impact of spasticity shown within this study might be explained by the higher GMFCS level (IV-V) of 'RaceRunning' athletes, of whom the majority were not able to walk (van der Linden et al. 2018).

Comparing participant characteristics of this sample, with previous studies predicting gait in children with CP (e.g. Balzer, Marsico et al. 2017), revealed that the included participants had lower levels of mobility and higher levels of spasticity (i.e. MAS total scores higher). Accordingly, in more severely impaired people with CP, the results of this study imply that, spasticity, in addition to impaired SVMC and muscle strength, limits mobility.

7.3.2 Considerations when testing reliability of a SVMC measure

As stated in chapter 2.3.3, reliability is defined as the degree to which a measurement is free from measurement error (Portney and Watkins 2000). The sources of measurement error are dependent on the measurement instrument applied as well as on the chosen population and other external factors.

Study 2 focused on the agreement within one rater and between two different raters. As the SCALE testing procedures are described in detail in the manual and because of the simplicity and ease of use of the SCALE, a well-standardized testing protocol helped to minimize possible error-sources. Furthermore, a second rating (by an independent rater) was achieved through video-recordings. Rating the SCALE from video recordings might be recommended, as it allows the sole concentration on scoring and not the simultaneous conduct of the assessment and communication with the participant. As the test duration of the SCALE is quite short (around 15 minutes), all participants could remain focused, concentrated and compliant. While intra-rater reliability (Kusumoto et al. 2016; Balzer et al. 2016) as well as inter-rater reliability (Kusumoto et al. 2016; Fowler et al. 2009) of the SCALE have been evaluated and shown to lie within excellent ranges, future studies might want to concentrate on test-retest reliability of this assessment tool. By testing this type of the SCALE's reliability, one could get an impression of the influence of potentially interfering factors, for instance how the participant's compliance or state of health might alter the testing conditions when applying the SCALE twice.

Considering absolute reliability of the SCALE, scores like the MDC changes were only reported in study 2. The two other studies on the SCALE's reliability presented relative reliability values. As absolute reliability (or measurement error) is critical for clinical as well as scientific work, future studies should also establish values of absolute reliability.

In relation to the SCALE-SI, the focus of the reliability testing lied less on the inter-rater reliability but more on testing the measurement error when applying the measurement tool twice (study 4). This type of reliability is known as test-retest reliability and recommended for all outcome measures. Due to the known variability of sEMG measures (i.e. electrode placement, impedance, muscle size), the expectations for between-session test-retest reliability were less, compared to the SCALE's excellent intra/inter reliability values (study 1). By strictly following the SEMINAN guidelines and by accurately noting down the distance of the electrode placement from bony landmarks, it was aimed to minimize error in relation to electrode placement. A customized seat was used to minimize the bias of movement artefacts upon the sEMG data. Movement artefacts of mainly the m. semitendinosus and m. gastrocnemius appeared less (testing within the pilot phase) when measurements took place on the seat. Nevertheless, movement artefacts appeared and are possible sources of error. Artefacts appeared most often during hip joint testing, during lifting and lowering the thigh from the seat, which sometimes caused a vibration on the sEMG sensor. Another possible source of bias might have been the setting of event markers (using a manual trigger) for the event windows for each of the three repetitive SCALE test-joint movements; i.e. definition of the start and endpoint of every single joint movement (e.g. start and end of knee extension and flexion movement one, two and three). In contrast to the original protocol of the SI (Voluntary Response Index), which uses a computer-generated audio signal as testing-starting and end point (Lim and Sherwood 2005), manual triggers simultaneous to the verbal command of the tester were applied in study 4. This technique might have been less precise and might have caused a greater variability in the time frame for each event window but was considered as the best possible alternative method to set event markers. Nevertheless, future studies might consider to either use the above mentioned method of a computer-generated audio signal or simultaneous recordings of the joint movements (e.g. goniometers) to obtain a more accurate way of establishing the start and end of the joint movements. Overall, the SCALE testing procedure seemed to be feasibly combined with the sEMG. The only exception here might have been the

testing of the SVMC of the hip joint in sitting position. Future studies might consider testing hip SVMC in accordance to the original SCALE protocol.

Thus, the lower ICC values at joint level might be explained by the aforementioned limitations. Nevertheless, relative reliability of the summed scores (but not the individual joint scores) is within clinically acceptable ranges ($ICC > 0.70$). Interpretation of absolute reliability values, i.e. whether a MDC 0.14 (SCALE-SI total score) is clinically acceptable, is not possible at the moment as real change estimates for SVMC of the lower extremity are not yet available.

7.3.3 Considerations concerning clinical utility of SVMC measures

Clinical utility is defined by the impact and usefulness of the results of a measure in relation to the individual, the family and society (Mokkink et al. 2016). Benefits and risks that have to be considered include the psychological, social and economic consequences of testing as well as the implications for health outcomes. Therefore, validity, reliability and responsiveness are critical measurement qualities of a tool, but if the instrument is not feasible within the certain practical context, it will be useless. Although clinical utility was not an explicit focus of this study, due to its importance for clinical as well as research practice, experiences with the SCALE and SCALE-SI will be discussed in this section. Furthermore, clinical utility of suggested measures will be discussed.

Application of the SCALE was feasible, as it can be performed within 15 minutes, does neither require any special training nor equipment, and is child-friendly. Compared to other SVMC assessments (SMC, m-Trost, Gillette's SMC), it assesses all five main lower extremity joints and applies a testing procedure and scoring system which aims to minimize the influence of muscle strength on SVMC. For those with less experience with the SCALE, it might be recommended to capture the assessment on film and to apply the scoring afterward using the video recording. As it is important that clinicians as well as researchers can readily and freely access the tools they require, the original SCALE as well as the German SCALE version are freely available to all. (http://uclaccp.org/images/ResearchPapers/SCALE%20Validity_appendix_%2009.pdf)

(http://www.kispi.uzh.ch/rza/de/forschende/publikationen/downloads/Documents/Balzer_Suppl.Material_SCALE_German.pdf)

The SCALE-SI testing items (i.e. joint movements) are, except for the hip, the same, but the testing procedure becomes more complex by the application of the sEMG. Measuring sEMG requires an sEMG system (approximately £50,000), specially trained testers (with hard- and software expertise as well as knowledge of the electrode placement) as well as additional time and knowledge to process and analyse the sEMG data. Because the correct electrode placement is critical, the duration of the SCALE-SI assessment is twice as long as the SCALE (30 min). Surface electrode placement (and removal) is not considered to be painful, however, in practice, removing electrodes was experienced as very uncomfortable by some children. The use of a plaster-remover (Niltact spray) helped to deal with this limitation. Because the application of the SCALE-SI demands for specific equipment and trained personnel, it might be feasible within some clinics as well as within the research setting but not within smaller practices, unlike the SCALE.

7.3.4 Considerations when testing responsiveness in SVMC measures

Responsiveness is “the ability of an instrument to detect change over time in the construct to be measured” (Mokkink et al. 2010). Accordingly, whenever a scientist or a clinician aims to detect, whether a patient’s health condition has changed over time due to recovery or an intervention, responsive outcome measures are required. Depending on the availability of a ‘gold standard’, responsiveness is tested either via criterion or construct validation.

In relation to the initial clinical question of this PhD, “Will my child learn to walk normally within your therapy?”, the desired outcome of this PhD would have been an evidence-based answer to this question. Right from the beginning of this project, it was clear that answering this question would only be possible with the existence and/or availability of a valid, reliable and responsive outcome measure. Only when using a tool with adequate psychometric properties, it is possible to carry out an intervention study on the effectiveness of a physiotherapy training strategy to improve SVMC of the lower extremity and subsequently gait quality in children with CP. In our rehabilitation centre, physiotherapeutically supporting the child to walk more ‘normally’ is mainly performed by training selective single-joint movements (e.g. ankle dorsiflexion) in static positions (e.g. in sitting, during standing, during initial contact position) as well as during functional movements, (e.g. during walking, via aids like FES). Accordingly, the systematic review aimed to evaluate psychometric properties

of outcome measures, measuring SVMC improvements of the lower extremities in children with UMN lesion. As the review results showed a relative scarcity of previously evaluated tools and the absence of responsiveness testing, the aim of this project was adapted and then geared toward developing a new SVMC outcome measure and testing its psychometric properties. Consequently, the SCALE-SI was developed, and its validity and reliability tested. However, as responsiveness testing of this new tool laid beyond the initial scope of this project, it remains the missing “step” in being able to address to the initial question of this project. Therefore, this section presents considerations and ideas in relation to a future responsiveness studies for the SCALE-SI. Firstly, considerations about finding an optimal intervention for testing responsiveness will be discussed by referring to the current debate about trainability of SVMC. Secondly, possible study designs for future responsiveness studies will be presented.

7.3.4.1 Finding an adequate intervention for responsiveness testing

In relation to SVC outcome measure responsiveness, an intervention is required, which is known to be effective in changing the outcome of interest in at least a part of the participant. Although this COSMIN definition appears to be sound, it gives rise to the question “which came first: the chicken or the egg”, as an intervention is needed that is known to improve SVMC. However, in order to establish the effectiveness of an intervention, a valid, reliable as well as responsive outcome measure is needed. It is this causality dilemma which exactly makes it so difficult to test responsiveness in outcome measures.

In relation to SVMC, this dilemma might even be exaggerated by the ongoing debate about trainability of SVMC in patients with UMN lesions. Mainly the following two issues are under current discussion. Firstly, there is an absence of consensus about if, to what extent or in what kind of patients impaired SVMC can be trained and how to do this most efficiently. Secondly, the additional value of training SVMC for the patients’ functioning in daily life is discussed.

Considering the first point of debate, up until now there is evidence missing as to what degree SVMC can be trained in patients suffering from a UMN lesion. Indeed, the impaired neuroanatomical structures cannot be restored or healed medically. Nevertheless, the concept of neuroplasticity as explained in section 2.2.9.1, describes the “ability of the brain to form and reorganize synaptic connections, especially in

response to learning or experience or following injury” (Nudo 2013). This activity or (use-dependent) plasticity of the brain (“neurons that fire together will wire together”) has been shown in animal, neurologically intact humans and in research with patients with UMN lesions. Although the majority of studies in this field have shown changes in upper extremity motor control, the fact that this evidence for SVMC of the lower extremity is currently lacking does not indicate that neuroplasticity does not exist for the lower extremities. It rather shows that measuring the neuroanatomical structures of the upper extremities is easier, as well as it shows that SVMC of the upper extremity is more commonly trained. While impaired SVMC is of great importance when carrying out fine motor movements of the hand, its accurate functioning is needed to a lesser extent for lower limb functional tasks like standing and walking, where muscle strength seems to be also of great importance for moving against gravity.

Due to the functional relation of many selective upper limb movements, SVMC training of the upper limbs is naturally imposed into therapy and the patient’s everyday life. In contrast, SVMC of the lower extremity, for instance by training selective ankle dorsiflexion, can be less well integrated into a purposeful activity training. Due to reduced possibilities of incorporating SVMC training of the lower extremity within functional training, the neuroplastic potential of the neuroanatomical structures innervating the lower extremity might have been underestimated. In this context, the lack of evidence concerning selective training of the lower extremity should have increased visibility and might help to guide future studies within this context.

The second main point of debate is the value of improved SVMC for the child’s quality of life. While some advocate to train solely on the activity level, aiming for maximizing functional performance of the patient, others argue that by training SVMC the patient’s movement control will improve and thereby his/her movement efficiency. This improved motor control potential would then lead to an uplift of the patient’s performance on the activity level. As an evidenced-based answer to this debate can only be given by the results of robust intervention studies, which falls beyond the scope of this PhD project, future research in this area is needed.

Furthermore, it becomes clear, that the selection of a reference measure to responsiveness of the SCALE-SI is currently based on clinical practice and not on scientific evidence. Based on i) the author’s clinical experience as a specialized neuro-paediatric physiotherapist as well as in relation to ii) the principles of neuroplasticity and due to iii) the results of the SVMC intervention studies (Chapter

2.2.8: Table 1), an intervention to assess the responsiveness of the SCALE-SI in children with CP would have to fulfil the following criteria. I) it should include functional activities; ii) be attractive / ('fun') and easy to understand (cognitive deficits). Further, it should iii) allow for a high number of repetitions, iv) be easy to standardise, v) and it should previously be shown to have an effect on SVMC. For example, an intervention using a robotic ankle/knee device in combination with virtual reality may fulfil these criteria. Although this intervention is not related to the patients' real daily activities, it would offer the possibility to playfully train selective ankle movements (Chen et al. 2016; Sukal-Moulton et al. 2014; Byanton et al. 2006). Furthermore, the use of virtual reality has shown to allow for standardization as well as for a high number of 'repetitions without repetitions' (Meyer-Heim and Van Hedel 2014). Alternatively, instead of using a robotic joint device, FES (to the dorsiflexors) may be a feasible intervention, both on its own (e.g. during single joint movement and functional movement as well as during activities like gait) or in combination with virtual reality. The additional benefit of FES would be that it acts directly on the muscles as well as that it adds sensory input to learned movements (Pool et al. 2014). In the opinion of the author, this would constitute two suitable intervention options for testing responsiveness of SVMC. Of course, interventions like NDT or Vojta also aim to improve SVMC, but as these interventions are less standardisable, they seem to be less suitable for responsiveness testing. Interventions like SDR or BTX, which are thought to influence SVMC indirectly by reducing the limiting effect of spasticity, might be alternative options.

7.3.4.2 Outlook study design for responsiveness testing of the SCALE-SI

The study design of a responsiveness study should control for as many as possible interfering factors. In order to be able to draw a robust conclusion regarding the ability of the SCALE-SI to measure a change in SVMC in children with CP, the following prerequisites should be fulfilled: i) an appropriated study design (approach) should be selected; ii) an adequate intervention, based on motor learning and neuroplasticity principles, should be applied (section 7.3.4.1), iii) a representative sample of children with spastic CP should be recruited; and iv) adequate measurement time points and comparison measures should be selected.

The aim of this section is to consider and explain the various aspects of a potential study protocol to establish responsiveness of the SCALE-SI in children with CP.

i) Study approach

Although the importance of responsiveness is well accepted, there is less consensus with regard to the way responsiveness should be assessed and reported. A review performed by Husted and colleagues in 2000 showed that there are two approaches known as “internal” and “external” responsiveness. Internal responsiveness measures responsiveness by the ability of the tool to record a change over a particular pre-specified time frame. Using this approach responsiveness is often quantified by effect sizes (type I-II).

Effect sizes have been introduced by Cohen to provide a standardized measure of the magnitude of an effect. These measures are used to interpret changes in health status or magnitudes of treatment effects (Cohen 1988). The COSMIN group, however, considers that an effect size only has a meaning as a measure of responsiveness if the magnitude of an effect (change) is known beforehand, because otherwise interpretation of the measured value cannot be clearly distinguished from the treatment effect (De Vet et al. 2011).

Therefore, they advocate the external responsiveness approach, in which responsiveness is quantified by the correlation between the change in the measure of interest and the corresponding change in a reference (gold standard) measure. However, other researchers question the meaningful application of quantifying responsiveness via correlation coefficients, as the clinical relevance of expected correlation thresholds is not as readily interpreted as effect size measures (Husted et al. 2000; Angst 2011).

This dilemma with regard to the assessment of responsiveness is particularly challenging in the context of this thesis, as in the field of SVMC neither prior established effect sizes exist, nor “gold standard” or “external reference (tool)” exist. For a future study assessing the responsiveness of measures of SVMC (like the SCALE-SI), the candidate therefore suggests quantifying responsiveness using the methods of the internal approach (effect sizes) and the external approach (correlation with a comparator instrument in absence of an external reference/gold standard measure).

ii) Change in SVMC through treatment or normal development

No matter which responsiveness approach is chosen, a change in the construct of interest, either due to (normal) development or treatment, must be observable.

With regard to development, changes in SVMC occur with maturation of the central nervous system. Although these changes are known to occur in children with CP, their timing and intensity vary greatly in this heterogeneous population. Further, although such a development approach might be interesting in relation to SVMC of the upper extremity, the author doubts its applicability to the lower extremity for the following reason. As previously explained, ambulatory children with CP might experience “learned-non-use” of the lower extremity associated with strict orthosis management regimes (common in Switzerland). Accordingly, natural development (maturation) of SVMC around the ankle, which may be primarily impaired due to their brain damage, might be further suppressed by the learned-non-use.

As the primary interest of this PhD is related to therapy-induced changes of SVMC, we would favour to apply an intervention to investigate the responsiveness of the SCALE-SI. The intervention of choice would be an activity-based motor learning approach including FES and motivating virtual reality gaming to increase therapy compliance of the younger patients. Although previous studies have shown that interventions such as robotic training or FES in combination with virtual reality (section 2.2.8) have an immediate effect on SVMC (Byanton et al. 2006), a longer, more intensive training programme may be necessary to ensure sufficient time for structural and functional changes in relation to SVMC to occur. Based on previous intervention studies (Chen et al. 2016; Sukal Moulton et al. 2014; Jung et al. 2013; Wu et al. 2011), an intervention duration of six weeks, with three training sessions per week, is suggested to induce and assess changes of SVMC in the lower extremity in children with CP.

An alternative approach to measure expected changes of SVMC might be found in interventions like SDR or BTX, which are thought to influence SVMC indirectly by reducing spasticity. Nevertheless, as this approach would be based on the controversial pathophysiological nature of SVMC, their expected treatment effect is more speculative than pre-established, and therefore considered to be less appropriate. Furthermore, as mentioned in chapter 2.2.4, (early) intervention aiming to decrease exaggerated afferent input (i.e. BTX) should be considered with caution. For these reasons, the author recommends therapeutic interventions which aim to increase activity in residual CST functioning instead.

iii) Sample

The population of interest would be, in accordance to COSMIN guidelines at least 30 children with CP, fulfilling the following inclusion criteria: clinical diagnosis of spastic CP, age between 5 and 18 years, GMFCS I-III, and ability to follow simple instructions. Children with a primarily dystonic or ataxic impairment, an unstable situation regarding their tonus-regulating medications and/or children who had a BTX injection within the last six months or any surgical correction within the last year of the lower extremity, should be excluded. Furthermore, only children, who are capable of understanding the training and assessment instructions, which is relevant concerning voluntary motor control, would be included. Regarding the GMFCS level, we would aim to recruit equally within the different levels to allow for inter-group comparison.

iv) Measurement procedure and tools

To evaluate internal responsiveness of the SCALE-SI, the SCALE-SI would be recorded before and after the intervention to calculate its effect size (Cohen's d or non-parametric effect size where appropriate). In order to quantify the SCALE-SI's responsiveness externally, the change in SCALE-SI score should be compared to the change on another SVMC measure score.

The author would suggest the SCALE assessment and Transcranial Magnetic Stimulation (TMS) as two possible comparator measures for external quantification of the responsiveness of the SCALE SI. As the SCALE's ability to detect changes might be limited in relation to its ordinal scoring system (sections 4.5.4; 6.5.4; 7.3.1.2), only moderate to low positive correlations between the changes scores of the SCALE-SI and SCALE are expected.

TMS is a non-invasive method measuring the integrity of the CST using sEMG recordings of the target muscle. A coil is placed on the head and creates a changing magnetic field over the motor cortex. This causes an electric current to flow from the cortex via electromagnetic induction to the related skeletal muscle. TMS has been used most extensively to evaluate physiological and impaired CST function (Beaulieu et al. 2014; Vry et al. 2008; Kesar et al. 2012). As stimulation of the muscles of the lower limb (e.g. m. tibialis anterior) is more difficult due to the neuroanatomical position of the foot on the motor cortex, the majority of studies are found on the upper extremity (Beaulieu, Massé-Alarie, et al. 2017). Nevertheless, one very interesting study in relation to the interest of this PhD was a study performed by Yang and

colleagues (Yang et al. 2013). They studied the probable effectiveness of early gait training in young children with hemiplegic CP (under the age of 2 years). One main objective of their study was to investigate the effectiveness of an intensive activity-based therapy program (on average 1 h/day, 4 days/week) on enhancing walking and gait quality in five hemiplegic children younger than 2 years. This very young age was selected on purpose in order to train these children in the critical period of CST maturation. Moreover, to prove the hypothesis that learning within this time window is more effective than after maturation has completed. Preliminary results show that improvements in walking (GMFM 66) and gait analysis outcomes, such as walking symmetry, exceeded previous reports, which supported their hypothesis. Their second goal was to refine TMS techniques for measuring motor and sensory pathways to and from the legs in a reliable manner in this young age group. By using several adjustments (e.g. use of a large, double-cone coil with an outside diameter of 125 mm, double pulses, and triggering in standing position to have a similar level of background activity while the child was playing), they concluded that this was possible, therewith allowing future large studies to measure the effectiveness of such early intervention program as well as in relation to its changes of CST functioning (e.g. motor evoked potential, stimulation threshold, latency, reflex response).

In accordance to such an adapted TMS measurement protocol (Yang et al. 2013), measuring Motor Evoked Potentials (MEP) of the tibialis anterior before and after the suggested intervention seems to be feasible. An increase of the amplitude of the MEP of the tibialis anterior would be associated with improved CST functioning and could be the comparator measure to compare changes in the SCALE-SI with. However, although that TMS measures the intactness of the most important pathway involved in SVMC, the CST misses several important characteristics of SVMC. Firstly, as activation of the muscle is stimulated externally, it does not require the participant's voluntary activation, which is central when referring back to the definition of SVMC (Sanger et al. 2006). Secondly, it only measures the motor output in relation to the stimulated area. Thereby, measurement and detection of involuntary movements (co- and mirror movements) is lacking. Furthermore, larger studies assessing the psychometric properties of TMS mainly focused on evaluating reliability of TMS in either neurological intact adults or adults with a neurological condition. A recent review on the consistency and measurement error of TMS in neurologically intact participants concluded that "evidence about the reliability of TMS outcomes is scarce and affected by several methodological and statistical problems" (Beaulieu, Flament, et al. 2017).

Accordingly, it would be hypothesized that only a moderate positive correlation between the change scores of the both measures would be found.

Finally, as an additional external and clinical anchor for evaluating the SCALE-SI's responsiveness, the therapist perception of the occurred change in SVMC could be assessed by asking the question "Do you think that SVMC has improved during the intervention?" and relating the answer (e.g. on a five point Likert scale) to the change recorded in SCALE-SI.

7.4 Common methodological considerations, limitations and future work

This section will cover further methodological considerations, limitations as well as ideas for future studies, which have not been covered in the previous sections of this chapter (7.3). Firstly, limitations in methodological considerations to the regression analyses of study 1 will be discussed. Secondly, common limitations of all studies in this thesis as well as ideas for future work will be presented.

7.4.1 Multicollinearity

One main methodological consideration in relation to the multiple regression analysis of study 1 (Chapter 3), was the problem of multicollinearity. Multicollinearity is present when the independent variables of a multiple regression analysis are strongly related to each other. This makes it impossible to obtain unique estimates of the explained variance as these variables account for the similar variance of the dependent variable. Thereby, their beta values are interchangeable, increasing the risk of type II error (Miles and Shevlin 2001). The correlation matrix in study 1, which aimed to predict gait capacity by lower extremity and trunk impairments, revealed high correlations ($p > 0.6$) between the MMT vs. SCALE, MMT vs. TCMS, and SCALE vs. TCMS. Furthermore, the average variance inflation factor of the starting model was above "1", which is considered as a threat to the validity of the model (Field 2005). Due to this multicollinearity, the SCALE and MMT scores explained a similar amount of mTUG variance and were removed from the final model on gait capacity, while the TCMS and MAS remained in the model. As already mentioned in study 1, these results do not indicate that SVMC and leg muscle strength does not influence gait capacity. The MAS, which on its own only correlated weakly with the mTUG, seems

to explain another part of the variance. Therefore, only the MAS and the variable with the highest beta value (TCMS) were kept in the final model.

This interpretation is supported by the results of the simple regression analyses, which showed that SCALE and MMT explained the second and the third largest amount of variance in mTUG (43% and 40%, respectively). A more robust statistical solution to deal with multicollinearity would be to run a factor analysis. How such a structural equation modelling could look like is presented in the study of Park and Lim (2013). Their study confirmed the construct of motor impairment and performed structural equation modelling between motor impairments (muscle strength, spasticity, range of motion, SVMC), gross motor function (Gross Motor Function Measure), and functional outcomes regarding activities of daily living (Functional Skills domain of the Paediatric Evaluation of Disability Inventory) in children with CP. Results of this study revealed that motor impairments had a large direct effect and explained the majority of the variance in gross motor function. Furthermore, the path coefficients of all motor impairments together as well as those of individual ones (SVMC; muscle strength) allowed the estimation of the influence of each factor on the dependent variable. Thereby, they were able to show that an increase of one standard deviation in motor impairments was associated with a decrease of 0.869 standard deviations in gross motor function. Further, they showed that an increase of one standard deviation in strength and SVMC was associated with an increase of 0.723 and 0.660 standard deviations in gross motor function, respectively. As this statistical approach requires a larger sample size (> 30-460), it might be considered for future studies investigating similar research questions within a larger sample (Wolf et al. 2015).

Furthermore, a future study investigating the influence of SVMC and other lower extremity and trunk control motor impairments by structural equation modelling might aim to increase the robustness of the measured values by using objective, pre-evaluated, measurement tools, instead of observer-dependent ones like in study 1. For instance, instead of applying the SCALE, the SCALE-SI could be used, or instead of using the MMT, an isokinetic measure should be applied. As outcome measures for the dependent variable, walking tests like the mTUG or other measures like three-dimensional gait analysis using a full body model (i.e. to measure lower and upper extremity as well as trunk kinematic, kinetics) or the 'Gait Profile Score' might be chosen. Thereby, in addition to temporal gait characteristics, kinematic data could be entered in the model as well. Possibly, results of such a study would allow for a more

in-depth understanding of the neurophysiological contribution of SVMC and the other selected variables on gait function in children with CP.

In conclusion, carrying out structural equation modelling would allow for the investigation of the influence of SVMC among other motor impairments on gait activities, solving the threat of multicollinearity.

7.4.2 Generalizability is limited to children with GMFCS levels I-III

A common methodological limitation of all studies in this thesis is that generalizability of their results is limited to children with CP with GMFCS levels I-III. The dominance of participants with a better level of walking ability (GMFCS I and II) should be considered when interpreting the results of the studies. Nevertheless, this restriction to children with higher gross motor abilities is related to the overall dominance of children with CP with severity levels GMFCS I and II in the entire population as well as to the cognitive demands of the SVMC testing procedure. As the latter requires that participants fully understand the purpose of the test-purpose (selective single-joint movement only, no co- and mirror movements), we had to exclude relatively more frequently participants with GMFCS levels III and IV compared to levels I and II. In general, it is more difficult to recruit children with lower mobility level and who have adequate cognitive functioning. Therefore, the underrepresentation of children with more severe mobility problems might also be present in future studies.

7.4.3 Estimation of clinical meaningful threshold

Another limitation of study 2 and 4 was the unequal sample distribution within the different GMFCS levels as well as the overall limited sample size (study 1: $n=39$; study 4: $n=24$). Based on sample size guidelines (Miles and Shevlin 2001), further statically procedures to estimate clinical meaningful thresholds in relation to the SCALE's / SCALE-SI's ability to discriminate between individual GMFCS levels were not performed. Future studies with larger samples might, therefore, consider performing a ROC curves to determine the SCALE-SI's ability to differentiate between adjacent GMFCS levels. SCALE-SI cut-off values for differentiating between GMFCS levels, will be determined at the level with the highest Youden Index ($= \text{sensitivity} + \text{specificity} - 1$), reflecting the best combination of sensitivity and specificity. Furthermore, the area under the curve can be calculated, which is a measure of the accuracy of the ROC analysis (Youden 1950). The area under the curve is given by a value between 0 and

1, whereas 0.5 represents chance and indicates results of a random process. Values between 0 and 0.5 indicate worse results than a random process, while a 1 indicates that there are no false positives or false negatives, i.e. the test is perfect. The research group around the Voluntary Response Index and the SI (Lee et al. 2004; McKay et al. 2004; McKay et al. 2005) computed a ROC analysis in patients with spinal cord injury. Their ROC results showed that the SI could not differentiate well between participants with higher motor functioning (American Spinal Injury Association levels D and E), as in these participants muscle activity was too close to that of healthy subjects, with a resulting clustering of SI values near 1.0 (ceiling effect).

Computing such analysis for the SCALE-SI could result in similar clinical meaningful findings as well as cut-off values.

7.4.4 Future work

Considering the mentioned limitation of the SCALE-SI algorithm being a combined quantification of all muscle activations, irrespectively if it is a matter of selective or pathological (co- and mirror-movements) activations, this results in the inability to say if mirror movements or co-movements are worse in the sense of SVMC. As this differentiation has until now not been scientifically investigated but might be relevant to improve efficient treatment planning in children with CP, one might consider upgrading the SI algorithm with an equation, which is capable of “weighing” non-physiological movements. For instance, by looking at the collected reference activation pattern of the neurological intact adults and children, and by applying methods like nonnegative matrix factorization (Steele et al. 2015), this might be accomplished.

Another important but until now un-investigated prerequisite of SVMC is adequate sensory input, i.e. proprioception. Considering the further establishment of the SCALE's or the SCALE-SI's convergent validity, future studies might aim to test the hypotheses that a deficit in lower extremity proprioception is correlated with a low level of SVMC. The biggest challenge within such a study might be to find an adequate measurement tool for proprioception, as measuring sensory deficits is challenging. Most recent scientific investigations as well as available measurement instruments testing sensory deficits in children with CP are designed to test somatosensory awareness of the upper extremity (e.g. Van de Winckel et al. 2013). One study investigating proprioception of the lower extremity used a custom-built device to

assess joint-position sense and kinaesthesia in the transverse plane of hip internal/external rotation. Their results showed that children with a diplegic or hemiplegic CP had proprioception deficits in all limbs (Wingert et al. 2009). Although their custom-made proprioception testing procedure seems to be adequate, evaluation of its psychometric properties is outstanding but essential if applied for the purpose of validating a SVMC tool.

Another topic of interest would be to investigate if SVMC has an impact on a child's participation and personal factors like self-esteem. In particular, self-esteem seems to be of interest in relation to SVMC as it is related to movement aesthetics. Due to the rhythmic pattern of walking, un-physiological walking patterns are readily detected and can cause social-stigmatization. This was also shown by one previous study (Riad et al. 2013). Riad and colleagues showed that the gait pattern deviating more from that of healthy peers (Gait Profile Score) of physically high functioning children with unilateral CP (GMFCS I-II) correlated with lower self-esteem. This was even more pronounced when non-physiological arm movement was involved.

Future studies might therefore investigate the impact of SVMC of the lower extremity in children with CP on gait (performances and capacity), self-esteem and its social consequences by using a multiple regression model or structural equational modelling.

7.5 Conclusion

The starting point of this PhD project was the author's desire to give parents from children with CP an evidenced-based answer to their frequently asked question "Will my child learn to walk normally within your therapy?". The origin of this question is mainly the fear that the child will be socially stigmatized due to an altered lower extremity control during gait. Besides this personal fear, research within the last decade has shown that impaired SVMC in the lower extremity can initiate a vicious cycle, limiting other body functions and hampering gross motor function.

Accordingly, the desired product of this PhD project was initially to provide an evidence-based answer to the question whether physiotherapy is effective in improving lower extremity SVMC in relation to normalizing the patient's gait pattern. Since an answer to this question implied an intervention study, a sensitive yet robust outcome measure for SVMC with established psychometric properties was needed.

As such a measure was missing, the initial aim of the project changed into the following intermediate goals:

- 1) finding new evidence in relation to the importance of SVMC for gait and its functional interconnection with other common motor impairments (study 1),
- 2) establishing assessments that could be applied in clinical routine examinations of SVMC in the German speaking medical field (by means of the German version of the SCALE) (study 2),
- 3) finding all available previously tested SVMC measurement instruments and evaluating their psychometric properties (study 3),
- 4) developing and testing a new objective SVMC outcome measure (SCALE-SI) and establish its psychometric properties (study 4),
- 5) providing suggestions for testing responsiveness of the SCALE-SI in a future study (Chapter 7: Final discussion).

Retrospectively, the initial question of this PhD project is still not answered. Nevertheless, by phrasing and realising necessary sub-goals, this thesis added important evidence and tools required for answering the question which is so frequently asked by parents in the future. As soon as future studies will have established responsiveness of the developed and tested SCALE-SI, an intervention study testing the effectiveness of physiotherapy in improving SVMC of the lower extremity, will provide the initially desired answer to the question “Will my child learn to walk normally within your therapy?”. Thereby, although the author could not reach her initial goal within the scope of this project, the process as well as all intermediate goals met along this journey were considered as being worthwhile in relation to their value for other scientists, clinicians and patients.

References

- AMMANN-REIFFER, C. BASTIAENEN, C.H., DE BIE, R.A. and VAN HEDEL, H.J.A., 2014. Measurement Properties of Gait-Related Outcomes in Youth With Neuromuscular Diagnoses: A Systematic Review. *Physical Therapy*. 94, pp.1067–1082.
- ANGST, F., 2011. The new COSMIN guidelines confront traditional concepts of responsiveness. *Medical Research Methodology* [online].11:152. [viewed 19 October 2018] Available from: doi: <https://doi.org/10.1186/1471-2288-11-152>.
- ARPIN, D.J., STUBERG, W., STERGIOU, N. and KURZ, M.J., 2013. Motor control of the lower extremity musculature in children with cerebral palsy. *Research in Developmental Disabilities*. 34(4), pp.1134–1143.
- ASLAN, S.C., CHOPRA, M.K., MCKAY, W.B., FOLZ, R.J. and OVECHKIN, A.V., 2013. Evaluation of Respiratory Muscle Activation Using Respiratory Motor Control Assessment (RMCA) in Individuals with Chronic Spinal Cord Injury. *Journal of Visualized Experiments* [online]. (77) [viewed 11 January 2016] Available from: <https://www.jove.com/video/50178/evaluation-respiratory-muscle-activation-using-respiratory-motor>.
- ASANUMA, H., LARSEN, K. AND YUMIYA, H., 1980. Peripheral input pathways to the monkey motor cortex. *Experimental Brain Research*. 38(3), pp.349–355.
- ATTIAS, M., BONNEFOY-MAZURE, A., LEMPEREUR, M., LASCOMBES, P., DE COULON, G. and ARMAND, S., 2014. Trunk movements during gait in cerebral palsy. *Clinical biomechanics (Bristol, Avon)*. 30(1), pp.28–32.
- AYALON, M., BEN-SIRA, D., HUTZLER, Y. and GILAD, T., 2000. Reliability of isokinetic strength measurements of the knee in children with cerebral palsy. *Developmental Medicine and Child Neurology*. (42), pp.398–402.
- BADAWI, N. and KEOGH, J.M., 2013. Causal pathways in cerebral palsy. *Journal of Paediatrics and Child Health*. 49(1), pp.5–8.
- BALZER, J., MARSICO, P., MITTEREGGER, E., VAN DER LINDEN, M.L., MERCER, T.H. and VAN HEDEL, H.J.A., 2016. Construct validity and reliability of the Selective Control Assessment of the Lower Extremity in children with cerebral palsy. *Developmental Medicine and Child Neurology*. 58(2), pp.167–172.
- BALZER, J., MARSICO, P., MARSICO, P., MITTEREGGER, E., VAN DER LINDEN, M.L., MERCER, T.H. and VAN HEDEL, H.J.A., 2017. Influence of trunk control and lower extremity impairments on gait capacity in children with cerebral palsy. *Disability and Rehabilitation* [online] 40(26), pp. 3164-3170. [viewed 01 May 2018] Available from: <https://doi.org/10.1080/09638288.2017.1380719>.

- BALZER, J., VAN DER LINDEN, M.L., MERCER, T.H. and VAN HEDEL, H.J.A., 2017. Selective voluntary motor control measures of the lower extremity in children with upper motor neuron lesions: a systematic review. *Developmental Medicine and Child Neurology*. 59(7), pp.699–705.
- BANDHOLM, T, ROSE, M.H., SLØK, R., SONNE-HOLM, S. and JENSEN, B.R., 2009. Ankle torque steadiness is related to muscle activation variability and coactivation in children with cerebral palsy. *Muscle and Nerve*, 40, pp.402–410.
- BANDHOLM, T., SLØK, R., SONNE-HOLM, S. and JENSEN, B.R, 2008. Assessment of torque-steadiness reliability at the ankle level in healthy young subjects: Implications for cerebral palsy. *European Journal of Applied Physiology*. 104, pp.609–615.
- BANDHOLM, T., JENSEN, B.R., NIELSEN, L.M., RASMUSSEN, H., BENCKE, J., CURTIS, D., PEDERSEN, S.A. and SONNE-HOLM, S., 2012. Neurorehabilitation with versus without resistance training after botulinum toxin treatment in children with cerebral palsy: A randomized pilot study. *Neurorehabilitation*. 30(4), pp.277–86.
- BARBER, C.E., 2008. A guide to physiotherapy in cerebral palsy. *Paediatrics and Child Health*. 18(9), pp.410–413.
- BARBER, L., HASTINGS-ISON, T., BAKER, R., BARRETT, R. and LICHTWARK, G., 2011. Medial gastrocnemius muscle volume and fascicle length in children aged 2 to 5 years with cerebral palsy. *Developmental Medicine and Child Neurology*. 53(6), pp.543–548.
- BARTHAS, F. and KWAN, A.C., 2017. Secondary Motor Cortex: Where “Sensory” Meets “Motor” in the Rodent Frontal Cortex Introduction: An Emerging View of M2 Function. *Trends in Neurosciences*. 40(3), pp.181–193.
- BARTHÉLEMY, D., GREY, M.J., NIELSEN, J.B. and BOUYER, L., 2011. Involvement of the corticospinal tract in the control of human gait. *Progress in Brain Research*. 192, pp.181–197.
- BAX, M., TYDEMAN, C. and FLODMARK, O., 2006. Clinical and MRI Correlates of Cerebral Palsy. *Jama*. 296(13), p.1602-1608.
- BAXTER, P., PANETH, N., LEVITON, A., GOLDSTEIN, M., BAX, M., DAMIANO, D., DAN, B. and JACOBSSON, B., 2008. The Definition and Classification of Cerebral Palsy. *Developmental Medicine and Child Neurology*. 49, pp.1–44.
- BEATON, D., BOMBARDIER, C., GUILLEMIN, F. and BOSI FERRAZ, M., 2007. Recommendations for the cross-cultural adaptation of the DASH and QuickDASH outcome measures. *American Academy of Orthopaedic Surgeons and Institute for Work and Health*. pp.1–45.

- BEAULIEU, L.D., FLAMAND, V., MASSÉ-ALARIE, H. and SCHNEIDER, C., 2017. Reliability and minimal detectable change of transcranial magnetic stimulation outcomes in healthy adults: A systematic review. *Brain Stimulation*. 10, pp.196–213.
- BEAULIEU, L.D., MASSÉ-ALARIE, H., RIBOT-CISCAR, E. and SCHNEIDER, C., 2017. Reliability of lower limb transcranial magnetic stimulation outcomes in the ipsi- and contralesional hemispheres of adults with chronic stroke. *Clinical Neurophysiology*. 128, p.1290–1298.
- BEAULIEU, L.D., MASSÉ-ALARIE, H., BROUWER, B. and SCHNEIDER, C., 2014. Brain control of volitional ankle tasks in people with chronic stroke and in healthy individuals. *Journal of the neurological sciences*. 338(1–2), pp.148–55.
- BEAULIEU, L.D. and SCHNEIDER, C., 2013. Effects of repetitive peripheral magnetic stimulation on normal or impaired motor control. A review. *Neurophysiologie clinique = Clinical neurophysiology*. 43(4), pp.251–60.
- BEGNOCHE, D.M., CHIARELLO, L.A., PALISANO, R.J., GRACEY, E.J., MCCOY, S.W. and ORLIN, M.N., 2016. Predictors of Independent Walking in Young Children With Cerebral Palsy. *Physical Therapy*. 96(2), pp.183–192.
- BHATTACHARYA, A. and LAHIRI, A., 2002. Mirror Movement in Clinical Practice. *Indian Academy of Clinical Medicine*. 3(2), pp.177–181.
- BISHOP, P.A. and HERRON, R.L., 2015. Use and Misuse of the Likert Item Responses and Other Ordinal Measures. *International journal of exercise science*. 8(3), pp.297–302.
- BOHANNON, R. and SMITH, M., 1987. Interrater reliability of a modified Ashworth scale of muscle spasticity. *Physical Therapy*,# 67, pp.206–207.
- BOJANIC, D., PETROVACKI-BALJ, BD JORGOVANOVIC, N. and ILIC, V., 2011. Quantification of dynamic EMG patterns during gait in children with cerebral palsy. *Journal of Neuroscience Methods*. 198(2), pp.325–331.
- BOWERMAN, B. and O'CONNEL, I R., 2000. *Linear Statistical Models: An Applied Approach*. 2nd ed., Belmont CA: Duxbury Classic.
- BOYD, R.N. and GRAHAM, H.K., 1999. Objective measurement of clinical findings in the use of botulinum toxin type A for the management of children with cerebral palsy. *European Journal of Neurology*. 6(suppl 4), pp.23–35.
- BRUCHEZ, R., JEQUIER GYGAX, M., ROCHES, S., FLUSS, J., JACQUIER, D., BALLABENI, P., GRUNT, S. and NEWMAN, C., 2016. Mirror therapy in children with hemiparesis: a randomized observer-blinded trial. *Developmental Medicine and Child Neurology*. 58(9), pp.970–978.
- BRUNNSTROMM, S., 1966. Motor testing procedures in hemiplegia: based on sequential recovery stages. *Physical Therapy*. 46, pp.357–375.

- BUCCINO, G., SOLODKIN, A. and SMALL, S.L., 2006. Functions of the mirror neuron system: implications for neurorehabilitation. *Cognitive and behavioral neurology: official journal of the Society for Behavioral and Cognitive Neurology*. 19(1), pp.55–63.
- BUCKON, C.E., THOMAS, S.S., HARRIS, G.E., PIATT, J.H., AIONA, M.D. and SUSSMAN, M.D., 2002. Objective measurement of muscle strength in children with spastic diplegia after selective dorsal rhizotomy. *Archives of Physical Medicine and Rehabilitation*. 83(4), pp.454–460.
- BURDEA, G., RABIN, B. and CHAPERON, A., 2011. Emotive , Cognitive and Motor Rehabilitation Post Severe Traumatic Brain Injury – A New Convergent Approach. In: *International Conference on Virtual Rehabilitation 2011, Zurich, Switzerland, 2011*.
- BURDEA, G.C., CIOI, D., KALE, A., JANES, W.E., ROSS, S.A. and ENGSBERG, J.R., 2013. Robotics and gaming to improve ankle strength, motor control, and function in children with cerebral palsy--a case study series. *IEEE transactions on neural systems and rehabilitation engineering: a publication of the IEEE Engineering in Medicine and Biology Society*. 21(2), pp.165–73.
- BYANTON, C., BOSSÉ, J., BRIEN, M., MCLEAN, J., MCCORMICK, A. and SVEISTRUP, H., 2006. Feasibility , Motivation , and Selective Motor Control : Virtual Reality Compared to Conventional Home Exercise in Children with Cerebral Palsy. *Cyberpsychology Behaviour*. 9(2), pp.123–128.
- CAEYENBERGHS, K., LEEMANS, A., GEURTS, M., TAYMANS, T., LINDEN, C.V., SMITS-ENGELSMAN, B.C., SUNAERT, S. and SWINNEN, S.P., 2010. Brain-behavior relationships in young traumatic brain injury patients: DTI metrics are highly correlated with postural control. *Human Brain Mapping*, 31(7), pp.992–1002.
- CAHILL-ROWLEY, K. and ROSE, J., 2014. Etiology of impaired selective motor control: emerging evidence and its implications for research and treatment in cerebral palsy. *Developmental Medicine and Child Neurology*. 56(6), pp.522–8.
- CAMPBELL, S.K., COLE, W., BOYNEWICZ, K., ZAWACKI, L.A., CLARK, A., GAEBLER-SPIRA, D., DEREGNIER, R.A., KURODA, M.M., KALE, D., BULANDA, M. and MADHAVAN, S., 2018. Behavior During Tethered Kicking in Infants with Periventricular Brain Injury. *Neurorehabilitation and Neural Repair*. 32(1), pp.62-72.
- CARR, L.J., HARRISON, L.M., EVANS, A.L. and STEPHENS, J.A., 1993. Patterns of central motor reorganization in hemiplegic cerebral palsy. *Brain*. 116(5), pp.1223–1247.
- CARR, L.J., 2014. Clinical Presentation. In: DAN, B., MAYSTON, M., PANETH, N. and ROSENBLOOM, L., eds. *Cerebral Palsy - Science and Clinical Practice*. London: Mac Keith Press, pp. 225–240.

- CARRARO, E., ZEME, S., TICCINELLI, V., MASSARONI, C., SANTIN, M., PERETTA, P., MARTINUZZI, A. and TREVISI, E., 2014. Multidimensional outcome measure of selective dorsal rhizotomy in spastic cerebral palsy. *European Journal of Paediatric Neurology*. 18(6), pp.704–713.
- CD-10-CM, 2018. 2018 nternational Statistical Classification of Diseases and Related Health Problems (ICD-10-CM) Diagnosis Code G80.9. Available from: <http://www.icd10data.com/ICD10CM/Codes/G00-G99/G80-G83/G80-/G80.9> [Accessed February 11, 2018].
- CHAPMAN, J.P., CHAPMAN, L.J. and ALLEN, J.J., 1987. The measurement of foot preference. *Neuropsychologia*. 25(3), pp.579–584.
- CHEN, K., WU, Y.N., REN, Y., LIU, L., GAEBLER-SPIRA, D., TANKARD, K., LEE, J., SONG, W., WANG, M. and ZHANG, L.Q., .2016. Home-Based Versus Laboratory-Based Robotic Ankle Training for Children With Cerebral Palsy: A Pilot Randomized Comparative Trial. *Archives of Physical Medicine and Rehabilitation*. 97(8), pp.1237–1243.
- CHRUSCIKOWSKI, E., FRY, N.R.D., NOBLE, J.J., GOUGH, M. and SHORTLAND, A.P., 2017. Selective motor control correlates with gait abnormality in children with cerebral palsy. *Gait and Posture*. 52, pp.107–109.
- CIOI, D., KaLE, A., BURDEA, G., ENGSBERG, J., JANES, W. and ROSS, S., 2011. Ankle control and strength training for children with cerebral palsy using the Rutgers Ankle CP: a case study. *IEEE International Conference on Rehabilitation Robotics* [online].10(30). [viewed 16 April 2016] Available from: <http://www.jneuroengrehab.com/content/10/1/30>, 2011.
- CLOWRY, G.J., 2007. The dependence of spinal cord development on corticospinal input and its significance in understanding and treating spastic cerebral palsy. *Neuroscience and Biobehavioral Reviews*, 31(8), pp.1114-1124.
- CURTIS, D.J., BUTLER, P., SAAVEDRA, S., BENCKE, J., KALLEMOSE, T., SONNE-HOLM, S. and WOOLLACOTT, M., 2014. The central role of trunk control in the gross motor function of children with cerebral palsy: a retrospective cross-sectional study. *Developmental Medicine and Child Neurology*. 57(4), pp.351-357.
- CUTHBERT, S. and GOODHEART, G., 2007. On the reliability and validity of manual muscle testing: a literature review. *Chiropractic and Osteopathy* [online].15(4). [viewed 16 April 2016] Available from [https:// doi: 10.1186/1746-1340-15-4](https://doi.org/10.1186/1746-1340-15-4).
- DAMIANO, D.L., MARTELOTTA, T.L., SULLIVAN, D.J., GRANATA, K.P. and ABEL, M.F., 2000. Muscle force production and functional performance in spastic cerebral palsy: relationship of cocontraction. *Archives of Physical Medicine and Rehabilitation*. 81(7), pp.895–900.

- DAMIANO, D.L., MARTELOTTA, T.L., QUINLIVAN, J.M. and ABEL, M.F., 2001. Deficits in eccentric versus concentric torque in children with spastic cerebral palsy. *Med Sci Sports Exerc.* 33, pp.117–122.
- DAMIANO, D.L., 2006. Activity, Activity, Activity: Rethinking Our Physical Therapy Approach to Cerebral Palsy. *Physical Therapy.* 86(11), pp.1534–1540.
- DAMIANO, D.L., PROSSER, L.A., CURATALO, L.A. and ALTER, K.E., 2012. Muscle Plasticity and Ankle Control After Repetitive Use of a Functional Electrical Stimulation Device for Foot Drop in Cerebral Palsy. *Neurorehabilitation and Neural Repair.* 27(3), pp.200–207.
- DAN, B., MAYSTON, M., PANETH, N. and ROSENBLOOM, L., eds. 2014. *Cerebral Palsy - Science and Clinical Practice*. London: Mac Keith Press.
- DEGELAEN, M., DE BORRE, L., KERCKHOFS, E., DE MEIRLEIR, L., BUYL, R., CHERON, G. and DAN, B., 2013. Influence of botulinum toxin therapy on postural control and lower limb intersegmental coordination in children with spastic cerebral palsy. *Toxins.* 5(1), pp.93–105.
- DESLOOVERE, K., MOLENAERS, G., FEYS, H., HUENAERTS, C., CALLEWAERT, B. and VAN DE WALLE, P., 2006. Do dynamic and static clinical measurements correlate with gait analysis parameters in children with cerebral palsy? *Gait and Posture.* 24(3), pp.302–313.
- DOBSON, F., 2010. Assessing selective motor control in children with cerebral palsy. *Developmental Medicine and Child Neurology.* 52(5), pp.409–10.
- DOORNIK, J., KUKKE, S., MCGILL, K., ROSE, J., SHERMAN-LEVINE, S. and SANGER, T.D., 2008. Oral Baclofen Increases Maximal Voluntary Neuromuscular Activation of Ankle Plantar Flexors in Children With Spasticity Due to Cerebral Palsy. *Journal of Child Neurology.* 23(6), pp.635–639.
- DOWNING, A.L., GANLEY, K.J., FAY, D.R. and ABBAS, J.J., 2009. Temporal characteristics of lower extremity moment generation in children with cerebral palsy. *Muscle and Nerve.* 39(6), pp.800–809.
- ENGSTBERG, J.R., ROSS, S.A., OLREE, K.S. and PARK, T.S., 2000. Ankle spasticity and strength in children with spastic diplegic cerebral palsy. *Developmental Medicine and Child Neurology.* 42, pp.42–47.
- ENGSTBERG, J.R., VAN EMMERIK, R.E., ROSS, S.A. and COLLINS, D.R., 2008. Use of relative phase as a measure of motor control at the ankle in persons with cerebral palsy: A preliminary study. *Journal of Applied Biomechanics.* 24(4), pp.382–390.
- ENGSTBERG, J.R., ROSS, S. and PARK, T.S., 2004. Quantifying active ankle range of motion in cerebral palsy following selective dorsal rhizotomy. *Journal of Applied Biomechanics.* 20, pp.103–111.

- EUNSON, P., 2016. Aetiology and epidemiology of cerebral palsy. *Paediatrics and Child Health (United Kingdom)*. 26(9), pp.367–372.
- EVERAERT, D.G., THOMPSON, A.K., CHONG, S.L. and STEIN, R.B., 2010. Does functional electrical stimulation for foot drop strengthen corticospinal connections? *Neurorehabilitation and Neural Repair*. 24(2), pp.168–177.
- EYRE, J.A., TAYLOR, J.P., VILLAGRA, F., SMITH, M. and MILLER, S., 2001. Evidence of activity-dependent withdrawal of corticospinal projections during human development. *Neurology*. 57(9), pp.1543-1554.
- FEDRIZZI, E., FACCHIN, P., MARZAROLI, M., PAGLIANO, E., BOTTEON, G., PERCIVALLE, L. and FAZZI, E., 2000. Predictors of Independent Walking in Children With Spastic Diplegia. *Journal of Child Neurology*. 15(4), pp.228–234.
- FERGUSON, J., 2017. The Who, What, and How of Health Outcome Measures. *Health Catalyst*. pp.1–7.
- FETTERS, L., CHEN, Y.P., JONSDOTTIR, J. and TRONICK, E.Z., 2004. Kicking coordination captures differences between full-term and premature infants with white matter disorder. *Human movement science*. 22(6), pp.729–48.
- FETTERS, L., SAPIR, I., CHEN, Y.P., KUBO, M. and TRONICK, E., 2010. Spontaneous kicking in full-term and preterm infants with and without white matter disorder. *Developmental psychobiology*. 52(6), pp.524–536.
- FIELD, A., 2005. *Discovering Statistics Using SPSS*. 2nd ed. London: SAGE.
- FLEUREN, J.F.M., VOERMAN, G.E., ERREN-WOLTERS, C.V., SNOEK, G.J., RIETMAN, J.S., HERMENS, H.J. and NENE, A.V., 2010. Stop using the Ashworth Scale for the assessment of spasticity. *Journal of Neurology, Neurosurgery and Psychiatry*. 81(1), pp.46–52.
- FORSSBERG, H., 2014. Neural basis of motor control. In: DAN, B., MAYSTON, M., PANETH, N. AND ROSENBLOOM, L., eds. *Cerebral Palsy - Science and Clinical Practice*. London: Mac Keith Press, pp. 199–224.
- FORSSBERG, H., 1985. Ontogeny of human locomotor control. 1. Infant stepping, supported locomotion and transition to independent locomotion. *Experimental Brain Research*. 57, pp.480–493.
- FOSANG, A.L., GALEA, M.P., MCCOY, A.T., REDDIHOUGH, D.S. and STORY, I., 2003. Measures of muscle and joint performance in the lower limb of children with cerebral palsy. *Developmental Medicine and Child Neurology*. 45(10), pp.664–670.
- FOWLER, E.G., 2010. Concepts in spasticity and selective motor control in children with spastic cerebral palsy. *Technology and Disability*. 22(4), pp 207-214.

- FOWLER, E.G., STAUDT, L.A., GREENBERG, M.B. and OPPENHEIM, W.L., 2009. Selective Control Assessment of the Lower Extremity (SCALE): development, validation, and interrater reliability of a clinical tool for patients with cerebral palsy. *Developmental Medicine and Child Neurology*. 51(8), pp.607–14.
- FOWLER, E.G. and GOLDBERG, E.J., 2009. The effect of lower extremity selective voluntary motor control on interjoint coordination during gait in children with spastic diplegic cerebral palsy. *Gait and Posture*. 29(1), pp.102–107.
- FOWLER, E.G., STAUDT, L. A. and GREENBERG, M.B., 2010. Lower-extremity selective voluntary motor control in patients with spastic cerebral palsy: Increased distal motor impairment. *Developmental Medicine and Child Neurology*. 52(3), pp.264–269.
- GAGE, J.R., Schwartz, M.H., Koop, S.E. and Novacheck, T.F., eds. 2009. *The Identification and Treatment of Gait Problems in Cerebral Palsy*. 2nd ed. New York: Wiley.
- GERBER, C.N., LABRUYÈRE, R.O.B. and VAN HEDEL, H.J., 2016. Reliability and Responsiveness of Upper Limb Motor Assessments for Children with Central Neuromotor Disorders: A Systematic Review. *Neurorehabilitation and Neural Repair*. 30(1), pp.19–39.
- GLADSTONE, D., DANELLS, C. and BLACK, S., 2002. The Fugl-Meyer assessment of motor recovery after stroke: a critical review of its measurement properties. *Neurorehabilitation and Neural Repair*. 16, pp.232–240.
- GLENN, O. A, LUDEMAN, N.A., BERMAN, J.I., WU, Y.W., LU, Y., BARTHA, A.I., VIGNERON, D.B., CHUNG, S.W., FERRIERO, D.M., BARKOVICH, A.J. and HENRY, R.G., 2007. Diffusion tensor MR imaging tractography of the pyramidal tracts correlates with clinical motor function in children with congenital hemiparesis. *American Journal of Neuroradiology*. 28, pp.1796–1802.
- GOLDBERG, E.J., REQUEJO, P.S. and FOWLER, E.G., 2011. Joint moment contributions to swing knee extension acceleration during gait in individuals with spastic diplegic cerebral palsy. *Gait and Posture*. 33(1), pp.66–70.
- GORDON, A., 2016. Impaired Voluntary Movement Control and Its Rehabilitation in Cerebral Palsy. *Adv Exp Med Biol*, 957, pp.291–311.
- HADZAGIC-CATIBUSIC, F., AVDAGIC, E., ZUBCEVIC, S. and UZICANIN, S. 2017. Brain Lesions in Children with Unilateral Spastic Cerebral Palsy. *Medical archives*. 71(1), pp.5–11.
- HALEY, S. and FRAGALA-PINKHAM, M., 2006. Interpreting change scores of tests and measures used in physical therapy. *Physical therapy*. 86(5), pp.735–743.

- HEATHCOCK, J.C., BHAT, A.N., LOBO, M.A. and GALLOWAY, J.C., 2004. The performance of infants born preterm and full-term in the mobile paradigm: Learning and memory. *Physical Therapy*. 84, pp.808–821.
- HEATHCOCK, J.C., BHAT, A.N., LOBO, M.A. and GALLOWAY, J.C., 2005. The relative kicking frequency of infants born full-term and preterm during learning and short-term and long-term memory periods of the mobile paradigm. *Physical Therapy*. 85, pp. 8–18.
- HUSTED, J.A., COOK, R.J., FAREWELL, V.T. and GLADMAN, D.D., 2000. Methods for assessing responsiveness: a critical review and recommendations. *Journal of Clinical Epidemiology*. 53(5), pp.459-468.
- VAN HEDEL, H.J.A., WIRTH, B. and CURT, A., 2010. Ankle motor skill is intact in spinal cord injury, unlike stroke: Implications for rehabilitation. *Neurology*. 74(16), pp.1271–1276.
- HEYRMAN, L., MOLENAERS, G., DESLOOVERE, K., VERHEYDEN, G., DE CAT, J., MONBALIU, E. and FEYS, H., 2011. A clinical tool to measure trunk control in children with cerebral palsy: the Trunk Control Measurement Scale. *Research in developmental disabilities*. 32(6), pp.2624–35.
- HEYRMAN, L., FEYS, H., MOLENAERS, G., JASPERS, E., MONARI, D., NIEUWENHUYIS, A. and DESLOOVERE, K., 2014. Altered trunk movements during gait in children with spastic diplegia: Compensatory or underlying trunk control deficit? *Research in Developmental Disabilities*. 35(9), pp.2044–2052.
- HICKS, T. and ONODERA, S., 2012. The mammalian red nucleus and its role in motor systems, including the emergence of bipedalism and language. *Progress in Neurobiology*. 96(2), pp.165-75.
- HINKLEY, L., WEBSTER, R.L., BYL, N.N. and NAGARAJAN, S.S., 2009. Neuroimaging characteristics of patients with focal hand dystonia. *Journal of Hand Therapy*. 22, pp.125–134.
- HOON, A.H., STASHINKO, E.E., NAGAE, L.M., LIN, D.D., KELLER, J., BASTIAN, A., CAMPBELL, M.L., LEVEY, E., MORI, S. and JOHNSTON, M.V., 2010. Sensory and motor deficits in children with cerebral palsy born preterm correlate with diffusion tensor imaging abnormalities in thalamocortical pathways. *Developmental Medicine and Child Neurology*. 51(9), pp.697–704.
- HUANG, C.Y., LIN, G.H., HUANG, Y.J., SONG, C.Y., LEE, Y.C., HOW, M.J., CHEN, Y.M., HSUEH, I.P., CHEN, M.H. and HSIEH, C.L., 2016. Improving the utility of the Brunnstrom recovery stages in patients with stroke: Validation and quantification. *Medicine (United States)* [online]. 95(31):e4508. Available from: [https://doi: 10.1097/MD.0000000000004508](https://doi.org/10.1097/MD.0000000000004508).

- HURVITZ, E. A., PETERSON, M. and FOWLER, E.G., 2014. Muscle tone, strength and movement disorders. In: DAN, B., MAYSTON, M., PANETH, N. AND ROSENBLOOM, L., eds. *Cerebral Palsy - Science and Clinical Practice*. London: Mac Keith Press, pp. 381–406.
- HUTCHENSON, G. and SOFRONIOU, N., 1999. *Multivariate Social Scientist - Introductory Statistics Using*, London: SAGE.
- JANG, S.H., 2014. The corticospinal tract from the viewpoint of brain rehabilitation. *Journal of Rehabilitation Medicine*. 46(31), pp.193–199.
- JOHNSTON, M., 2009. Plasticity in the developing brain: implications for rehabilitation. *Developmental Disabilities Research Reviews*. 15(9), pp.94–101.
- JØRGENSEN, L., PALUDAN-MÜLLER, A.S., LAURSEN, D.R.T., SAVOVIĆ, S., BOUTRON, I., STERNE, J.A.C., HIGGINS, J.P.T. and HRÓBJARTSSON, A., 2016. Evaluation of the Cochrane tool for assessing risk of bias in randomized clinical trials: overview of published comments and analysis of user practice in Cochrane and non-Cochrane reviews. *Systematic Reviews*, 0(5), pp.80.
- JUNG, J.W., HER, J.G. and KO, J., 2013. Effect of Strength Training of Ankle Plantarflexors on Selective Voluntary Motor Control, Gait Parameters, and Gross Motor Function of Children with Cerebral Palsy. *Journal of Physical Therapy Science*. 25(10), pp.1259–1263.
- KARABAY, İ., ÖZTÜRK, G.T., MALAS, F.Ü., KARA, M., TIFTİK, T., ERSÖZ, M. and ÖZÇAKAR, L., 2015. Short-Term Effects of Neuromuscular Electrical Stimulation on Muscle Architecture of the Tibialis Anterior and Gastrocnemius in Children with Cerebral Palsy. *American Journal of Physical Medicine and Rehabilitation*. 94(9), pp.728–733.
- KATZ, R., ROVAI, G. and BRAIT, C., 1992. Objective quantification of spastic hypertonia: correlation with clinical findings. *Archives of Physical Medicine and Rehabilitation*. 73, pp.339–347.
- KENDALL, F., MCCREARY, E. and PROVANCE, P., 1993. *Muscles: Testing and Function*., Baltimore ;Williams and Wilkins.
- KESAR, T.M., SAWAKI, L., BURDETTE, J.H., CABRERA, M.N., KOLASKI, K., SMITH, B.P., O'SHEA, T.M., KOMAN, L.A. and WITTENBERG, G.F., 2012. Motor cortical functional geometry in cerebral palsy and its relationship to disability. *Clinical Neurophysiology*. 123(7), pp.1383–1390.
- KIM, W.H. AND PARK, E.Y., 2011. Causal relation between spasticity, strength, gross motor function, and functional outcome in children with cerebral palsy: a path analysis. *Developmental Medicine and Child Neurology*. 53, pp.68–73.

- KIYAMA, M.T., FUKUDOME, K., HIYOSHI, T., UMEMOTO, A., YOSHIMOTO, Y. and MAEDA, T., 2011. The Loss of Dexterity in the Bilateral Lower Extremities in Patients with Stroke. *Journal of Applied Biomechanics*. 27(2), pp.122-129.
- KNUTSON, J., JAYME, S. and CHAE, J., 2011. A Novel Neuromuscular Electrical Stimulation Treatment for Recovery of Ankle Dorsiflexion in Chronic Hemiplegia: *Am J Phys Med Rehabil*. 89(8), pp.672–682.
- KOERTE, I., EFTIMOV, L., LAUBENDER, R.P., ESSLINGER, O., SCHROEDER, A.S., ERTL-WAGNER, B., WAHLLAENDER-DANEK, U., HEINEN, F. and DANEK, A., 2010. Mirror movements in healthy humans across the lifespan: effects of development and ageing. *Developmental Medicine and Child Neurology*. 52(12), pp.1106-1112.
- KOO, T.K. and LI, M.Y., 2016. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *Journal of Chiropractic Medicine*. 15(2), pp.155–163.
- KUO, H.C., FRIEL, K.M. and GORDON, A.M., 2018. Neurophysiological mechanisms and functional impact of mirror movements in children with unilateral spastic cerebral palsy. *Developmental Medicine and Child Neurology*. 60(2), pp:155-161.
- KUSUMOTO, Y., TAKAKI, K., MATSUDA, T. and NITTA, O., 2016. Relation of selective voluntary motor control of the lower extremity and extensor strength of the knee joint in children with spastic diplegia. *Journal of physical therapy science*. 28(6), pp.1868–1871.
- KUSUMOTO, Y., HANAO, M., TAKAKI, K., MATSUDA, T. and NITTA, O., 2016. Reliability and validity of the japanese version of the selective control assessment of the lower extremity tool among patients with spastic cerebral palsy. *Journal of Physical Therapy Science*. 28(12), pp.3316–3319.
- KWON, H.G., SON, S.M. and JANG, S.H., 2014. Development of the transcallosal motor fiber from the corticospinal tract in the human brain: diffusion tensor imaging study. *Frontiers in Human Neuroscience* [online]. 8:153. Available from: <https://doi.org/10.3389/fnhum.2014.00153>
- LAUER, R.T., STACKHOUSE, C., SHEWOKIS, P.A., SMITH, B.T., ORLIN, M. and MCCARTHY, J.J., 2005. Assessment of wavelet analysis of gait in children with typical development and cerebral palsy. *Journal of Biomechanics*. 38, pp.1351–1357.
- LAUER, R.T., STACKHOUSE, C.A., SHEWOKIS, P.A., SMITH, B.T., TUCKER, C.A. and MCCARTHY, J., 2007. A time – frequency based electromyographic analysis technique for use in cerebral palsy. *Gait and Posture*. 26, pp.420–427.

- LAWRENCE, D.G. and KUYPERS, H.G. 1968. The functional organization of the motor system in the monkey. II. The effects of lesions of the descending brain-stem pathways. *Brain*. 91(1), pp.15-36.
- LEBEL, C., GEE, M., CAMICOLI, R., WIELER, M., MARTIN, W. and BEAULIEU, C., 2012. Diffusion tensor imaging of white matter tract evolution over the lifespan. *NeuroImage*. 60(1), pp.340–352.
- LEBIEDOWSKA, M.K. and FISK, J.R., 2003. Quantitative evaluation of reflex and voluntary activity in children with spasticity. *Archives of Physical Medicine and Rehabilitation*. 84(6), pp.828–837.
- LEBIEDOWSKA, M.K., GAEBLER-SPIRA, D., BURNS, R.S. and FISK, J.R., 2004. Biomechanic characteristics of patients with spastic and dystonic hypertonia in cerebral palsy. *Archives of Physical Medicine and Rehabilitation*, 85(6), pp.875–880.
- LEE, D.C., LIM, H.K., MCKAY, W.B., PRIEBE, M.M., HOLMES, S.A. and SHERWOOD, A.M., 2004. Toward an objective interpretation of surface EMG patterns: a voluntary response index (VRI). *Journal of electromyography and kinesiology: official journal of the International Society of Electrophysiological Kinesiology*. 14(3), pp.379–88.
- LEE, J.D., PARK, H.J., PARK, E.S., OH, M.K., PARK, B., RHA, D.W., CHO, S.R., KIM, E.Y., PARK, J.Y., KIM, C.H., KIM, D.G. and PARK, C.I., 2011. Motor pathway injury in patients with periventricular leucomalacia and spastic diplegia. *Brain*. 134, pp.1199–1210.
- LENNON, N., THORPE, D., BALEMANS, A.C., FRAGALA-PINKHAM, M., O'NEIL, M., BJORNSEN, K., BOYD, R. and DALLMEIJER, A.J., 2015. Research in Developmental Disabilities Review article The clinimetric properties of aerobic and anaerobic fitness measures in adults with cerebral palsy: A systematic review of the literature. *Research in Developmental Disabilities*. 45(46), pp.316–328.
- LEVIN, M.F., DESROSIERS, J., BEAUCHEMIN, D., BERGERON, N. and ROCHETTE, A., 2004. Development and validation of a scale for rating motor compensations used for reaching in patients with hemiparesis: the reaching performance scale. *Physical therapy*. 84(1), pp.8–22.
- LEVIN, M.F., KLEIM, J.A. and WOLF, S.L., 2009. What do motor “recovery” and “compensation” mean in patients following stroke? *Neurorehabilitation and neural repair*. 23(4), pp.313–9.
- LEVIN, M.F., 2010. Associations between sensorimotor impairments and motor ability deficits: the pieces of the puzzle are beginning to fall into place. *Developmental Medicine and Child Neurology*. 52(9), p.788.

- LEVIN, M.F. and PANTURIN, E., 2011. Sensorimotor Integration for Functional Recovery and the Bobath Approach. *Motor Control*. 15(2), pp.285–301.
- LIM, H., 2015. Correlation between the selective control assessment of lower extremity and pediatric balance scale scores in children with spastic cerebral palsy. *Journal of physical therapy science*. 27(12), pp.3645–9.
- LIM, H.K., LEE, D.C., MCKAY, W.B., PRIEBE, M.M., HOLMES, S.A. and SHERWOOD, A.M. 2005. Neurophysiological assessment of lower-limb voluntary control in incomplete spinal cord injury. *Spinal cord*. 43, pp.283–90.
- LIM, H.K. and SHERWOOD, A.M., 2005. Reliability of surface electromyographic measurements from subjects with spinal cord injury during voluntary motor tasks. *The Journal of Rehabilitation Research and Development*. 42(4), p.413.
- VAN DER LINDEN, M.L., JAHED, S., TENNANT, N. and VERHEUL, M.H.G., 2018. The influence of lower limb impairments on RaceRunning performance in athletes with hypertonia, ataxia or athetosis. *Gait and Posture*. 61, pp.362–367.
- LOTZE, M., 2003. Motor learning elicited by voluntary drive. *Brain*. 126(4), pp.866–872.
- LOVETT, R. and MARTIN, E., 1916. Certain aspects of infantile paralysis with a description of a method of muscle testing. *Jama*. 66, pp.729–733.
- LÖWING, K. and CARLBERG, E.B., 2009. Reliability of the selective motor control scale in children with cerebral palsy. *Advances in Physiotherapy*. 11(2), pp.58–63.
- LUDEMAN, N. A., BERMAN, J.I., WU, Y.W., JEREMY, R.J., KORNAK, J., BARTHA, A.I., BARKOVICH, A.J., FERRIERO, D.M., HENRY, R.G. and GLENN, O.A., 2008. Diffusion tensor imaging of the pyramidal tracts in infants with motor dysfunction. *Neurology*. 71, pp.1676–1682.
- MANIKOWSKA, F., CHEN, B.P., JÓŹWIAK, M. and LEBIEDOWSKA, M.K., 2016. Assessment of selective motor control in clinical Gillette's test using electromyograph. *European journal of physical and rehabilitation medicine*. 52(2), pp.176–185.
- MANTOVANI, J. and SCRUTTON, D., 2014. Historical Perspective. In: DAN, B., MAYSTON, M., PANETH, N. and ROSENBLOOM, L., eds. *Cerebral Palsy - Science and Clinical Practice*. London: Mac Keith Press, pp. 3–15.
- MARSICO, P., FRONTZEK-WEPS, V., BALZER, J. and VAN HEDEL, H.J.A., 2017. Hypertonia Assessment Tool: Reliability and Validity in Children with Neuromotor Disorders. *Journal of Child Neurology*. 32(1), pp.132-138.

- MARSICO, P., Mitteregger, E., Balzer, J. and H.J.A., 2017. The Trunk Control Measurement Scale: reliability and discriminative validity in children and young people with neuromotor disorders. *Developmental Medicine and Child Neurology*. 59(7), pp.706-712.
- MARTIN, J.H., CHOY, M., PULLMAN, S. and MENG, Z., 2004. Corticospinal system development depends on motor experience. *Journal of Neuroscience*. 24, pp.2122–2132.
- MAYSTON, M.J., 2001. People with cerebral palsy: effects of and perspectives for therapy. *Neural plasticity*. 8(1–2), pp.51–69.
- MAYSTON, M.J., 2014. Intervention planning, implementation and evaluation. In: DAN, B., MAYSTON, M., PANETH, N. and ROSENBLUM, L., eds. *Cerebral Palsy - Science and Clinical Practice*. London: Mac Keith Press, pp. 329–360.
- MCKAY, W.B., LIM HK, PRIEBE, M.M., STOKIC, D.S. and SHERWOOD, A.M., 2004. Clinical neurophysiological assessment of residual motor control in post-spinal cord injury paralysis. *Neurorehabilitation and neural repair*. 18(3), pp.144–53..
- MCKAY, W.B., LEE, D.C., LIM, H.K., HOLMES, S.A. and SHERWOOD, A.M., 2005. Neurophysiological examination of the corticospinal system and voluntary motor control in motor-incomplete human spinal cord injury. *Experimental brain research*. 163(3), pp.379–87..
- MEYER-HEIM, A. and VAN HEDEL, H.J.A., 2014. Roboterunterstützte und computerbasierte Neurorehabilitation für Kinder: Die Geschichte dahinter. *Praxis*. 103(15), pp.883–892.
- MEYNS, P., VAN GESTEL, L., BRUIJN, S.M., DESLOOVERE, K., SWINNEN, S.P. and DUYSSENS, J., 2012. Is interlimb coordination during walking preserved in children with cerebral palsy? *Research in Developmental Disabilities*. 33(5), pp.1418–1428.
- MILES, J. and SHEVLIN, M., 2001. *Applying Regression and Correlation. A Guide for Students and Researchers.*, London. SAGE Publications.
- MILLER, F., ed. 2007. *Physical Therapy of Cerebral Palsy*. New York: Springer.
- MITTEREGGER, E., MARSICO, P., BALZER, J. and VAN HEDEL, H.J.A., 2015. Translation and construct validity of the Trunk Control Measurement Scale in children and youths with brain lesions. *Research in Developmental Disabilities* [online]. 45–46. [viewed 7 March 2016] Available from: <https://doi:10.1016/j.ridd.2015.08.007>.
- MOKKINK, L.B. and TERWEE, C.B., 2010. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. *Quality of Life Research*. 19(4), pp. 539-49.

- MOKKINK, L.B., PRINSEN, C.A., BOUTER, L.M., VET, H.C. and TERWEE, C.B., 2016. The Consensus-based standards for the selection of health measurement INstruments (COSMIN) and how to select an outcome measurement instrument. *Brazilian Journal of Physical Therapy*. 20(2),pp.105-13.
- MÖLLER, F., ULMER, S., WOLFF, S., STEPHANI, U. and JANSEN, O., 2005. Cortical reorganization in children with connatal spastic hemiparesis -- a functional magnetic resonance imaging (fMRI) study. *Rofo*. 177(11), p.1552–1561.
- MUTLU, A., LIVANELIOGLU, A. and GUNEL, M.K., 2008. Reliability of Ashworth and Modified Ashworth scales in children with spastic cerebral palsy. *BMC musculoskeletal disorders* [online]. 9:44. [viewed 6 October 2014] Available from: <https://doi.org/10.1186/1471-2474-9-44>.
- NAGHDI, S., EBRAHIMI, I., ASGARI, A., OLYAEI, G.R., KAZEMNEJAD, A., MANSOURI, K. and ANSARI, N.N., 2007. A preliminary study into the criterion validity of the Modified Modified Ashworth Scale using the new measure of the alpha motoneuron excitability in spastic hemiplegia. *Electromyography and Clinical Neurophysiology*. 47, pp.187–192.
- NATHAN, P.W., SMITH, M.C. and DEACON, P., 1990. The corticospinal tracts in man. Course and location of fibres at different segmental levels. *Brain*. 113(2), pp.303-324.
- NOBLE, J.J., CHARLES-EDWARDS, G.D., KEEVIL, S.F., LEWIS, A.P., GOUGH, M., and SHORTLAND, A.P., 2014. Intramuscular fat in ambulant young adults with bilateral spastic cerebral palsy. *BMC Musculoskeletal Disorders*.15(1), pp.1–8.
- NOBLE, J.J., FRY, N.R., LEWIS, A.P., KEEVIL, S.F., GOUGH, M. and SHORTLAND, A.P., 2014. Lower limb muscle volumes in bilateral spastic cerebral palsy. *Brain AND Development*. 36(4), pp.294–300.
- NOBLE, J.J., CHRUSCIKOWSKI, E., FRY, N.R.D., LEWIS, A.P., GOUGH, M., and SHORTLAND, A.P., 2017. The relationship between lower limb muscle volume and body mass in ambulant individuals with bilateral cerebral palsy. *BMC Neurology*. 17(1), pp.1–9.
- NORMAN, G., 2010. Likert scales, levels of measurement and the “laws” of statistics. *Advances in Health Sciences Education*. 15(5), pp.625–632.
- NUDO, R.J., 2013. Recovery after brain injury: mechanisms and principles. *Human Neuroscience* [online] 7. [viewed 23 July 2016] Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3870954/> pp.1–14.
- OLREE, K.S., ENGSBERG, J.R., ROSS, S.A., and PARK, T.S., 2000. Changes in synergistic movement patterns after selective dorsal rhizotomy. *Developmental Medicine and Child Neurology*. 42(5), pp.297–303.

- OLSEN, J.E., ROSS, S.A., FOREMAN, M.H. and ENGSBERG, J.R., 2013. Changes in muscle activation following ankle strength training in children with spastic cerebral palsy: an electromyography feasibility case report. *Physical & occupational therapy in pediatrics*. 33(2), pp.230–42.
- OMRANI, M., KAUFMAN, M.T., HATSOPOULOS, N.G., and CHENEY, P.D., 2017. Perspectives on classical controversies about the motor cortex. *Journal of Neurophysiology*. 18(3), pp.1828-1848.
- OSKOUI, M., COUTINHO, F., DYKEMAN, J., JETTÉ, N. and PRINGSHEIM, T., 2013. An update on the prevalence of cerebral palsy: A systematic review and meta-analysis. *Developmental Medicine and Child Neurology*. 55(6), pp.509–519.
- OSTENSJØ, S., CARLBERG, E.B., and VØLLESTAD, N.K., 2004. Motor impairments in young children with cerebral palsy : relationship to gross motor function and everyday activities. *Developmental Medicine and Child Neurology*. 46(9): pp.580–589.
- PALISANO, R., SNIDER, L. and ORLIN, M., 2004. Recent advances in physical and occupational therapy for children with cerebral palsy. *Seminars in Pediatric Neurology*. 11(1), pp.66–77.
- PANDIAN, S., ARYA, K.N. and DAVIDSON, E.W.R., 2012. Comparison of Brunnstrom movement therapy and Motor Relearning Program in rehabilitation of post-stroke hemiparetic hand: a randomized trial. *Journal of bodywork and movement therapies*. 16(3), pp.330–7.
- PARK, E.Y. and KIM, W.H., 2013. Structural equation modeling of motor impairment, gross motor function, and the functional outcome in children with cerebral palsy. *Research in developmental disabilities*. 34(5), pp.1731–9.
- PARK, E., BAEK, S. and PARK, S., 2016. Systematic review of the effects of mirror therapy in children with cerebral palsy. *Journal of Physical Therapy Science*. 28(11), pp.3227–3231.
- PEACOCK, W.J., 2009. The neuronal control of movement. In J. R. Gage et al., eds. *The Identification and Treatment of Gait Problems in Cerebral Palsy*. New York: Wiley.
- PEIPER, A., 1963. *Cerebral Function in Infancy and Childhood*. Consultant., New York.
- PERRY, J.E., 1993. Determinants of muscle function in the spastic lower extremity. *Clin Orthop Relat Res*. 288, p.10–26.
- PHILIPPS, C.G., 1975. Laying the Ghost of 'Muscles Versus Movements'. *Canadian Journal of Neurological Science*, 2(3), pp.209-218.

- PHILLIPS, C.G. and PORTER, R., 1977. Corticospinal Neurones: Their Role in Movement. *Monographs of the Physiological Society* [online]. 34(v-xii). [viewed 11 October 2018] Available from: <https://physoc.onlinelibrary.wiley.com/doi/pdf/10.1113/expphysiol.1978.sp002442>
- PIERCE, S.R., LAUER, R.T., SHEWOKIS, P.A., RUBERTONE, J.A. and ORLIN, M.N., 2006. Test-Retest Reliability of Isokinetic Dynamometry for the Assessment of Spasticity of the Knee Flexors and Knee Extensors in Children With Cerebral Palsy. *Archives of Physical Medicine and Rehabilitation*. 87(5), pp.697–702.
- POOL, D., BLACKMORE, A.M., BEAR, N. and VALENTINE, J., 2014. Effects of short-term daily community walk aide use on children with unilateral spastic cerebral palsy. *Pediatric Physical Therapy*. 26(3), pp.308–317.
- PORTER, R. and LEMON, R., eds. 1993. *Corticospinal Function and Voluntary Movement*. Oxford: Oxford University Press.
- PORTNEY, L.G. and WATKINS, M.P., 2000. *Foundations of clinical research - Application to practice*. 2nd ed. London: Prentice-Hall.
- POSTANS, N.J. and GRANAT, M.H., 1999. Correction of Gait Abnormalities in Cerebral Palsy Children using Functional Electrical Stimulation. *Gait and Posture*. 10(1):pp.87.
- PROSSER, L.A., LEE, S.C., BARBE, M.F., VANSANT, A.F. and LAUER, R.T., 2010. Trunk and hip muscle activity in early walkers with and without cerebral palsy--a frequency analysis. *Journal of electromyography and kinesiology: official journal of the International Society of Electrophysiological Kinesiology*. 20(5), pp.851–859.
- PROSSER, L.A., CURATALO, L.A., ALTER, K.E. and DAMIANO, D.L., 2012. Acceptability and potential effectiveness of a foot drop stimulator in children and adolescents with cerebral palsy. *Developmental Medicine and Child Neurology*. 54(11), pp.1044–1049.
- RATHELOT, J.A. and STRICK, P.L., 2006. Muscle representation in the macaque motor cortex: An anatomical perspective. *Proceedings of the National Academy of Sciences*. 103(21), pp.8257–8262.
- RHA, D.W., CHANG, W.H., KIM, J., SIM, E.G. and PARK, E.S., 2012. Comparing quantitative tractography metrics of motor and sensory pathways in children with periventricular leukomalacia and different levels of gross motor function. *Neuroradiology*, 54(6), pp.615–621.
- RHA, D.W., CAHILL-ROWLEY, K., YOUNG, J., TORBURN, L., STEPHENSON, K., and ROSE, J., 2015. Biomechanical and clinical correlates of swing-phase knee flexion in individuals with spastic cerebral palsy who walk with flexed-knee gait. *Archives of Physical Medicine and Rehabilitation*. 96(3), pp.511–517.

- RHA, D.W., CAHILL-ROWLEY, K., YOUNG, J., TORBURN, L., STEPHENSON, K., and ROSE, J., 2016. Biomechanical and Clinical Correlates of Stance-Phase Knee Flexion in Persons With Spastic Cerebral Palsy. *PM and R*. 8(1), pp.11–18.
- RIAD, J., BROSTRÖM, E. and LANGIUS-EKLÖF, A., 2013. Do movement deviations influence self-esteem and sense of coherence in mild unilateral cerebral palsy? *Journal of Pediatric Orthopaedics*. 33(3), pp.298–302.
- RIOS, D.C., GILBERTSON, T., MCCOY, S.W., PRICE, R., GUTMAN, K., MILLER, K.E., FECHKO, A. and MORITZ, C.T., 2013. NeuroGame Therapy to improve wrist control in children with cerebral palsy: a case series. *Developmental neurorehabilitation*. 16(6), pp.398–409.
- ROSE, J. and MCGILL, K.C., 2005. Neuromuscular activation and motor-unit firing characteristics in cerebral palsy. *Developmental Medicine and Child Neurology*. 47(Mvc), pp.329–336.
- ROSENBAUM, P. and STEWART, D., 2004. The World Health Organization International Classification of Functioning, Disability, and Health: A Model to Guide Clinical Thinking, Practice and Research in the Field of Cerebral Palsy. *Seminars in Pediatric Neurology*. 11(1), pp.5–10.
- ROSENBAUM, P.L., 2014. Definition and Clinical Classification. In: DAN, B., MAYSTON, M., PANETH, N. and ROSENBLOOM, L., eds. *Cerebral Palsy - Science and Clinical Practice*. London: Mac Keith Press, pp. 17–27.
- ROSENTHAL, R., 1994. Parametric measures of effect size. In: COOPER, H. and HEDGES, L.V. eds. *The handbook of research synthesis*. New York: Russell Sage Foundation, pp. 231-244.
- ROSS, S. and ENGSBERG, J.R., 2002. Relation between spasticity and strength in individuals with spastic diplegic cerebral palsy. *Developmental Medicine and Child Neurology*. 44, pp.148–157.
- ROSS, S. and ENGSBERG, J.R., 2007. Relationships Between Spasticity, Strength, Gait, and the GMFM-66 in Persons With Spastic Diplegia Cerebral Palsy. *Archives of Physical Medicine and Rehabilitation*. 88, pp.1114–1120.
- ROTHWELL, J., 1987. *Control of human voluntary movement*, London: Croom Helm.
- RUAN, J., BLUDAU, S., PALOMERO-GALLAGHER, N., CASPERS, S., MOHLBERG, H., EICKHOFF, S.B., SEITZ, R.J. and AMUNTS, K., 2018. Cytoarchitecture, probability maps, and functions of the human supplementary and pre-supplementary motor areas. *Brain Structure and Function*. 223(9), pp.4169–4186.

- SÆTHER, R., HELBOSTAD, J.L., ADDE, L., BRÆNDVIK, S., LYDERSEN, S. and VIK, T., 2014. The relationship between trunk control in sitting and during gait in children and adolescents with cerebral palsy. *Developmental Med Child Neurology*. 57(4), pp.344-350.
- SANGER, T.D., DELGADO, M.R., GAEBLER-SPIRA, D., HALLETT, M., and MINK, J.W., 2003. Classification and Definition of Disorders Causing Hypertonia in Childhood. *Pediatrics*. 111(1), pp.e89–e97.
- SANGER, T.D., CHEN, D., DELGADO, M.R., GAEBLER-SPIRA, D., HALLETT, M., and MINK, J.W., 2006. Definition and classification of negative motor signs in childhood. *Pediatrics*. 118(5), pp.2159–2167.
- SCHIARITI, V. and MASSE, L.C., 2014. Relevant Areas of Functioning in Children With Cerebral Palsy Based on the International Classification of Functioning, Disability and Health Coding System: A Clinical Perspective. *Journal of Child Neurology*. 30(2), pp.216–222.
- SCPE, 2000. Surveillance of cerebral palsy in Europe: a collaboration of cerebral palsy surveys and registers. Surveillance of Cerebral Palsy in Europe (SCPE). *Developmental Medicine and Child Neurology*. 42 (12), pp.816-824.
- SCPE, 2013. *Decision tree for Cerebral Palsy Look at Classification Tree*. [online]. [viewed 11 March 2014] Available from: <http://www.scpenetwork.eu/assets/Uploads/SCPE-CP-Decision-tree.pdf>.
- SHANG, Q., MA, C.Y., LV, N., LV, Z.L., YAN, Y.B., WU, Z.R., LI, J.J., DUAN, J.L. and ZHU, C.L., 2015. Clinical study of cerebral palsy in children with periventricular leukomalacia. *Experimental and Therapeutic Medicine*. 9(4), pp.1336–1344.
- SHISHOV, N., MELZER, I. and BAR-HAIM, S., 2017. Parameters and Measures in Assessment of Motor Learning in Neurorehabilitation; A Systematic Review of the Literature. *Frontiers in Human Neuroscience* [online]. 24,11:28. [viewed 20 December 2017] Available from: <https://doi:10.3389/fnhum.2017.00082>
- SHORTLAND, A.P., 2011. Muscle volume and motor development in spastic cerebral palsy. *Developmental Medicine and Child Neurology* [online]. 53(6). [viewed 5 April 2015] Available from: <https://doi:10.1111/j.1469-8749.2011.03926.x>.
- SHORTLAND, A.P., 2017. European Society of Movement Analysis for Adults and Children (ESMAC) Gait Course. *Muscle requirements for gait*. [presentation]. Trondheim: 6-9 September.
- SHUMAN, B.R., SCHWARTZ, M.H. and STEELE, K.M., 2017. Electromyography Data Processing Impacts Muscle Synergies during Gait for Unimpaired Children and Children with Cerebral Palsy. *Frontiers in Computational Neuroscience* [online]. 11: 50. [viewed 7 December 2017] Available from: <https://doi:10.3389/fncom.2017.00050>

- SMALL, S.L., BUCCINO, G. and SOLODKIN, A., 2012. The mirror neuron system and treatment of stroke. *Developmental Psychobiology*. 54, pp.293–310.
- SMITH, L.R., LEE, K.S., WARD, S.R., CHAMBERS, H.G., and LIEBER, R.L., 2011. Hamstring contractures in children with spastic cerebral palsy result from a stiffer extracellular matrix and increased in vivo sarcomere length. *Journal of Physiology*. 589(10), pp.2625–2639.
- SMITH, L.R., CHAMBERS, H. and LIEBER, R.L., 2013. Reduced satellite cell population may lead to contractures in children with cerebral palsy. *Developmental Medicine and Child Neurology*. 31(9), pp.1713–1723.
- SMITH, M., NOTGHI, L. and DAN, B., 2014. Clinical Neurophysiology. In: DAN, B., MAYSTON, M., PANETH, N. and ROSENBLOOM, L., eds. *Cerebral Palsy - Science and Clinical Practice*. London: Mac Keith Press, pp. 279–284.
- SMITS, D.W., VAN GROENESTIJN, A.C., KETELAAR, M., SCHOLTES, V.A., BECHER J.G. and GORTER, J.W., 2010. Selective motor control of the lower extremities in children with cerebral palsy: inter-rater reliability of two tests. *Developmental neurorehabilitation*. 13(4), pp.258–265.
- STAUDT, M., PAVLOVA, M., BÖHM, S., GRODD, W. and KRÄGELOH-MANN, I., 2003. Pyramidal tract damage correlates with motor dysfunction in bilateral periventricular leukomalacia (PVL). *Neuropediatrics*. 34(4), pp.182–8.
- STAUDT, M., 2007. (Re-)organization of the developing human brain following periventricular white matter lesions. *Neuroscience and Biobehavioral Reviews*. 31(8), pp.1150–1156.
- SUKAL-MOULTON, T., CLANCY, T., ZHANG, L.Q. and GAEBLER-SPIRA, D., 2014. Clinical application of a robotic ankle training program for cerebral palsy compared to the research laboratory application: does it translate to practice? *Archives of Physical Medicine and Rehabilitation*. 95(8), pp.1433–40.
- SÜTBEYAZ, S., YAVUZER, G., SEZER, N. and KOSEOGLU, B.F., 2007. Mirror Therapy Enhances Lower-Extremity Motor Recovery and Motor Functioning After Stroke: A Randomized Controlled Trial. *Archives of Physical Medicine and Rehabilitation*. 88(May), pp.555–559.
- TEDROFF, K., KNUTSON, L.M. and SODERBERG, G.L., 2006. Synergistic muscle activation during maximum voluntary contractions in children with and without spastic cerebral palsy. *Developmental Medicine and Child Neurology*. 48(10), pp.789–96.
- TERWEE, C.B., BOT, S.D., DE BOER, M.R., VAN DER WINDT, D.A., KNOL, D.L., DEKKER, J., BOUTER, L.M. and DE VET, H.C.W., 2007. Quality criteria were proposed for measurement properties of health status questionnaires. *Journal of Clinical Epidemiology*. 60, pp.34–42.

- TERWEE, C.B., JANSMA, E.P., RIPHAGEN, I.I. and DE VET, H.C.W., 2009. Development of a methodological PubMed search filter for finding studies on measurement properties of measurement instruments. *Quality of Life Research*. 18, pp.1115–1123.
- THOMAS, B., EYSEN, M., PEETERS, R., MOLENAERS, G., VAN HECKE, P., DE COCK, P. and SUNAERT, S., 2005. Quantitative diffusion tensor imaging in cerebral palsy due to periventricular white matter injury. *Brain*, 128, pp.2562–2577.
- TROJAN, S. and POKORNÝ, J., 1999. Theoretical aspects of neuroplasticity. *Physiological Research*. 48(2), pp.87–97.
- VAN TULDER, M., FURLAN, A., BOMBARDIER, C. and BOUTER, L., 2003. Editorial Board of the Cochrane Collaboration Back Review Group. Updated method guidelines for systematic reviews in the cochrane collaboration back review group. *Spine*, 28, pp.1290–1299.
- VANMECHELEN, I.M., SHORTLAND, A.P. and NOBLE, J.J., 2018. Lower limb muscle volume estimation from maximum cross-sectional area and muscle length in cerebral palsy and typically developing individuals. *Clinical Biomechanics*. 51, pp.40–44.
- VELLEMAN, P.F. and WILKINSON, L., 1993. Nominal , Ordinal , Interval , and Ratio Typologies are Misleading. *The American Statistician*. 47(1), pp.1–19.
- DE VET, H.C.W., TERWEE, C.B., MOKKINK, L.B., and KNOL, D.L., 2011. *Measurement in Medicine* 1st ed. H. C. W. de Vet et al., eds., New York: Cambridge University Press.
- VOORMAN, J.M., DALLMEIJER, A.J., KNOL, D.L., LANKHORST, G.J. and BECHER, J.G., 2007. Prospective longitudinal study of gross motor function in children with cerebral palsy. *Archives of Physical Medicine and Rehabilitation*. 88(7), pp.871–876.
- VOS, R.C., BECHER, J.G., VOORMAN, J.M., GORTER, J.W., VAN ECK, M., VAN MEETEREN, J., SMITS, D.W., TWISK J.W. and DALLMEIJER, A.J., 2016. Longitudinal association between gross motor capacity and neuromusculoskeletal function in children and youth with cerebral palsy. *Archives of Physical Medicine and Rehabilitation*. 97(8), pp.1329–37.
- VRY, J., LINDER-LUCHT, M., BERWECK, S., BONATI, U., HODAPP, M., UHL, M., FAIST, M. and MALL, V., 2008. Altered cortical inhibitory function in children with spastic diplegia: A TMS study. *Experimental Brain Research*. 186, pp.611–618.
- WILLIAMS, E.N., CARROLL, S.G., REDDIHOUGH, D.S., PHILLIPS, B.A. and GALEA M.P., 2005. Investigation of the timed “Up & Go” test in children. *Developmental Medicine and Child Neurology*. 47(8), pp.518–524.

- VAN DE WINCKEL, A., VERHEYDEN, G., WENDEROTH, N., PEETERS, R., SUNAERT, S., VAN HECKE, W., DE COCK, P., DESLOOVERE, K., EYSEN, M. and FEYS H., 2013. Does somatosensory discrimination activate different brain areas in children with unilateral cerebral palsy compared to typically developing children? An fMRI study. *Research in Developmental Disabilities*. 34(5), pp.1710–1720.
- WELNIARZ, Q., DUSART, I. and ROZE, E., 2017. The corticospinal tract: Evolution, development, and human disorders. *Developmental Neurobiology*, 77(7), pp.810-829.
- WINGERT, J.R., BURTON, H., SINCLAIR, R.J., BRUNSTROM, J.E. and DAMIANO, D.L., 2009. Joint-Position Sense and Kinesthesia in Cerebral Palsy Jason. *Archives of Physical Medicine and Rehabilitation*. 90(3), pp.447–453.
- WITTENBERG, G.F., 2009. Motor mapping in cerebral palsy. *Developmental Medicine and Child Neurology*. 51 Suppl 4, pp.134–139.
- WOLF, E.J., HARRINGTON, K.M., CLARK, S.L. and MILLER, M.W., 2015. Sample Size Requirements for Structural Equation Models: An Evaluation of Power, Bias, and Solution Propriety. *National Institutes of Health*. 76(6), pp.913–934.
- WOLF, S.L., CATLIN, P.A., ELLIS, M., ARCHER, A.L., MORGAN, B. and PIACENTINO, A., 2001. Stroke Assessing Wolf Motor Function Test as Outcome Measure for Research in Patients After Assessing Wolf Motor Function Test as Outcome Measure for Research in Patients After Stroke Subjects and Methods. *Stroke*. 32, pp.1635–1639.
- WOODBURY, M.L., VELOZO, C.A., RICHARDS, L.G., DUNCAN, P.W., STUDENSKI, S. and LAI, S.M., 2007. Dimensionality and construct validity of the Fugl-Meyer Assessment of the upper extremity. *Archives of Physical Medicine and Rehabilitation*. 88(6), pp.715–23.
- WRIGHT, F. and MAJNEMER, A., 2014. The Concept of a Toolbox of Outcome Measures for Children With Cerebral Palsy. *Journal of Child Neurology*. 29(8), pp.1055–1065.
- YANG, H.S., KWON, H.G., HONG, J.H., HONG, C.P. and JANG, S.H., 2011. The rubrospinal tract in the human brain: Diffusion tensor imaging study. *Neuroscience Letters*. 54(1), pp.45–48.
- YANG, J.F., LIVINGSTONE, D., BRUNTON, K., KIM, D., LOPETINSKY, B., ROY, F., ZEWDIE, E., PATRICK, S.K., ANDERSEN, J., KIRTON, A., WATT, J.M., YAGER, J. and GORASSINI, M., 2013. Training to enhance walking in children with cerebral palsy: Are we missing the window of opportunity? *Seminars in Pediatric Neurology*. 20(2), pp.106–115.

- YAO, J., CHEN, A., CARMONA, C. and DEWALD, JP., 2009. Cortical overlap of joint representations contributes to the loss of independent joint control following stroke. *Neuroimage*. 45, p.490–499.
- YEO, S.S., CHANG, M.C., KWON, Y.H., JUNG, Y.J. and JANG, S.H., 2012. Corticoreticular pathway in the human brain: Diffusion tensor tractography study. *Neuroscience Letters*. 508(1), pp.9–12.
- YOUDEM, W.J., 1950. Index for Rating Diagnostic Tests. *Cancer*, 3(1), pp.32–35.
- ZOGHI, M., GALEA, M. and MORGAN, D., 2013. A Brain Motor Control Assessment (BMCA) protocol for upper limb function. *PloS one* [online]. 8(11):e79483. viewed 13 April 2016] Available from: <https://doi.org/10.1371/journal.pone.0079483>
- ZWAAN, E., BECHER, J.G. and HARLAAR, J., 2012. Synergy of EMG patterns in gait as an objective measure of muscle selectivity in children with spastic cerebral palsy. *Gait and Posture*. 35(1), pp.111–115.

Appendices

Appendix 1: COSMIN checklist.....	187
Appendix 2: Search-Construct SVMC measurement-tool list	210
Appendix 3: Overview of SVMC measurement-tools.....	212
Appendix 4: Study - German Version SCALE	214
Appendix 5: Study 2 - English SCALE version (original)	220
Appendix 6: Study 3 - Search terms for all databases.....	224
Appendix 7: Study 3 - established specified COSMIN rating rules	230
Appendix 8: Study 3 - Quality Criteria for Measurement Properties	232
Appendix 9: Study 3 - Evidence Level of Measurement Properties	233
Appendix 10: Publication study 1	234
Appendix 11: Publication study 2.....	242
Appendix 12: Publication study 2.....	249

Appendix 1: COSMIN checklist

COSMIN checklist with 4-point scale

Contact

CB Terwee, PhD
VU University Medical Center
Department of Epidemiology and Biostatistics
EMGO Institute for Health and Care Research
1081 BT Amsterdam
The Netherlands
Website: www.cosmin.nl, www.emgo.nl
E-mail: cb.terwee@vumc.nl



Instructions

This version of the COSMIN checklist is recommended for use in systematic reviews of measurement properties. With this version it is possible to calculate overall methodological quality scores per study on a measurement property. A methodological quality score per box is obtained by taking the lowest rating of any item in a box ('worse score counts'). For example, if for a reliability study one item in the box 'Reliability' is scored poor, the methodological quality of that reliability study is rated as poor. The Interpretability box and the Generalizability box are mainly used as data extraction forms. We recommend to use the Interpretability box to extract all information on the interpretability issues described in this box (e.g. norm scores, floor-ceiling effects, minimal important change) of the instruments under study from the included articles. Similar, we recommend to use the Generalizability box to extract data on the characteristics of the study population and sampling procedure. Therefore no scoring system was developed for these boxes.

This scoring system is described in this paper:

Terwee CB, Mokkink LB, Knol DL, Ostelo RWJG, Bouter LM, de Vet HCW. Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist. *Quality of Life Research* 2011, July 6 [epub ahead of print].

Step 1. Evaluated measurement properties in the article

	Internal consistency	Box A
	Reliability	Box B
	Measurement error	Box C
	Content validity	Box D
	Structural validity	Box E
	Hypotheses testing	Box F
	Cross-cultural validity	Box G
	Criterion validity	Box H
	Responsiveness	Box I

Step 2. Determining if the statistical method used in the article are based on CTT or IRT

Box General requirements for studies that applied Item Response Theory (IRT) models

	excellent	good	fair	poor
<p>Was the IRT model used adequately described? e.g. One Parameter Logistic 1 Model</p> <p>(OPLM), Partial Credit Model (PCM), Graded Response Model (GRM)</p>	IRT model adequately described	IRT model not adequately described		
<p>Was the computer software package used adequately described? e.g. 2 RUMM2020,</p> <p>WINSTEPS, OPLM, MULTILOG, PARSCALE, BILOG, NLMIXED</p>	Software package adequately described	Software package not adequately described		
<p>3 Was the method of estimation used adequately described? e.g. conditional maximum likelihood (CML), marginal maximum likelihood (MML)</p>	Method of estimation adequately described	Method of estimation not adequately described		
<p>Were the assumptions for estimating parameters of the IRT model checked? 4 e.g. unidimensionality, local independence, and item fit (e.g. differential item functioning (DIF))</p>	assumptions of the IRT model checked	assumptions of the IRT model partly checked	assumptions of the IRT model not checked or unknown	

To obtain a total score for the methodological quality of studies that use IRT methods, the 'worse score counts' algorithm should be applied to the IRT box in combination with the box of the measurement property that was evaluated in the IRT study. For example, if IRT methods are used to study internal consistency and item 4 in the IRT box is scored fair, while the items in the internal consistency box (box A) are all scored as good or excellent, the methodological quality score for internal consistency will be fair. However, if any of the items in box A is scored poor, the methodological quality score for internal consistency will be poor.

Step 3. Determining if a study meets the standards for good methodological quality

Box A. Internal consistency				
	excellent	good	fair	poor
1 Does the scale consist of effect indicators, i.e. is it based on a reflective model? <i>Design requirements</i>				
2 Was the percentage of missing items given?	Percentage of missing items described	Percentage of missing items NOT described		
3 Was there a description of how missing items were handled?	Described how missing items were handled	Not described but it can be deduced how missing items were handled	Not clear how missing items were handled	
4 Was the sample size included in the internal consistency analysis adequate?	Adequate sample size (≥ 100)	Good sample size (50-99)	Moderate sample size (30-49)	Small sample size (< 30)
5 Was the unidimensionality of the scale checked? i.e. was factor analysis or IRT model applied?	Factor analysis performed in the study population	Authors refer to another study in which factor analysis was performed in a similar study population	Authors refer to another study in which factor analysis was performed, but not in a similar study population	Factor analysis NOT performed and no reference to another study
6 Was the sample size included in the unidimensionality analysis adequate?	7* #items and ≥ 100	5* #items and ≥ 100 OR 6-7* #items but < 100	5* #items but < 100	< 5 * #items

7	Was an internal consistency statistic calculated for each (unidimensional) (sub)scale separately?	Internal consistency statistic calculated for each subscale separately			Internal consistency statistic NOT calculated for each subscale separately
8	Were there any important flaws in the design or methods of the study?	No other important methodological flaws in the design or execution of the study		Other minor methodological flaws in the design or execution of the study	Other important methodological flaws in the design or execution of the study
<i>Statistical methods</i>					
9	for Classical Test Theory (CTT), continuous scores: Was Cronbach's alpha calculated?	Cronbach's alpha calculated		Only item-total correlations calculated	No Cronbach's alpha and no item-total correlations calculated
10	for CTT, dichotomous scores: Was Cronbach's alpha or KR-20 calculated?	Cronbach's alpha or KR-20 calculated		Only item-total correlations calculated	No Cronbach's alpha or KR-20 and no item-total correlations calculated
11	for IRT: Was a goodness of fit statistic at a global level calculated? E.g. χ^2 , reliability coefficient of estimated latent trait value (index of (subject or item) separation)	Goodness of fit statistic at a global level calculated			Goodness of fit statistic at a global level NOT calculated

NB. Item 1 is used to determine whether internal consistency is relevant for the instrument under study. It is not used to rate the quality of the study.

Box B. Reliability: relative measures (including test-retest reliability, inter-rater reliability and intra-rater reliability)					
		excellent	good	fair	poor
Design requirements					
1	Was the percentage of missing items given?	Percentage of missing items described	Percentage of missing items NOT described		
2	Was there a description of how missing items were handled?	Described how missing items were handled	Not described but it can be deduced how missing items were handled	Not clear how missing items were handled	
3	Was the sample size included in the analysis adequate?	Adequate sample size (≥100)	Good sample size (50-99)	Moderate sample size (30-49)	Small sample size (<30)
4	Were at least two measurements available?	At least two measurements			Only one measurement
5	Were the administrations independent?	Independent measurements	Assumable that the measurements were independent	Doubtful whether the measurements were independent	measurements NOT independent
6	Was the time interval stated?	Time interval stated		Time interval NOT stated	
7	Were patients stable in the interim period on the construct to be measured?	Patients were stable (evidence provided)	Assumable that patients were stable	Unclear if patients were stable	Patients were NOT stable
8	Was the time interval appropriate?	Time interval appropriate		Doubtful whether time interval was appropriate	Time interval NOT appropriate

9	Were the test conditions similar for both measurements? e.g. type of administration, environment, instructions	Test conditions were similar (evidence provided)	Assumable that test conditions were similar	Unclear if test conditions were similar	Test conditions were NOT similar
10	Were there any important flaws in the design or methods of the study?	No other important methodological flaws in the design or execution of the study		Other minor methodological flaws in the design or execution of the study	Other important methodological flaws in the design or execution of the study
<i>Statistical methods</i>					
11	for continuous scores: Was an intraclass correlation coefficient (ICC) calculated?	ICC calculated and model or formula of the ICC is described	ICC calculated but model or formula of the ICC not described or not optimal. Pearson or Spearman correlation coefficient calculated with evidence provided that no systematic change has occurred	Pearson or Spearman correlation coefficient calculated WITHOUT evidence provided that no systematic change has occurred or WITH evidence that systematic change has occurred	No ICC or Pearson or Spearman correlations calculated
12	for dichotomous/nominal/ordinal scores: Was kappa calculated?	Kappa calculated			Only percentage agreement calculated
13	for ordinal scores: Was a weighted kappa calculated?	Weighted Kappa calculated		Unweighted Kappa calculated	Only percentage agreement calculated
14	for ordinal scores: Was the weighting scheme described? e.g. linear, quadratic	Weighting scheme described	Weighting scheme NOT described		

Box C. Measurement error: absolute measures					
		excellent	good	fair	poor
Design requirements					
1	Was the percentage of missing items given?	Percentage of missing items described	Percentage of missing items NOT described		
2	Was there a description of how missing items were handled?	Described how missing items were handled	Not described but it can be deduced how missing items were handled	Not clear how missing items were handled	
3	Was the sample size included in the analysis adequate?	Adequate sample size (≥100)	Good sample size (50-99)	Moderate sample size (30-49)	Small sample size (<30)
4	Were at least two measurements available?	At least two measurements			Only one measurement
5	Were the administrations independent?	Independent measurements	Assumable that the measurements were independent	Doubtful whether the measurements were independent	measurements NOT independent
6	Was the time interval stated?	Time interval stated		Time interval NOT stated	
7	Were patients stable in the interim period on the construct to be measured?	Patients were stable (evidence provided)	Assumable that patients were stable	Unclear if patients were stable	Patients were NOT stable
8	Was the time interval appropriate?	Time interval appropriate		Doubtful whether time interval was appropriate	Time interval NOT appropriate

9	Were the test conditions similar for both measurements? e.g. type of administration, environment, instructions	Test conditions were similar (evidence provided)	Assumable that test conditions were similar	Unclear if test conditions were similar	Test conditions were NOT similar
10	Were there any important flaws in the design or methods of the study?	No other important methodological flaws in the design or execution of the study		Other minor methodological flaws in the design or execution of the study	Other important methodological flaws in the design or execution of the study
<i>Statistical methods</i>					
11	for CTT: Was the Standard Error of Measurement (SEM), Smallest Detectable Change (SDC) or Limits of Agreement (LoA) calculated?	SEM, SDC, or LoA calculated	Possible to calculate LoA from the data presented		SEM calculated based on Cronbach's alpha, or on SD from another population

Box D. Content validity (including face validity)					
		excellent	good	fair	poor
General requirements					
1	Was there an assessment of whether all items refer to relevant aspects of the construct to be measured?	Assessed if all items refer to relevant aspects of the construct to be measured		Aspects of the construct to be measured poorly described AND this was not taken into consideration	NOT assessed if all items refer to relevant aspects of the construct to be measured

2	Was there an assessment of whether all items are relevant for the study population? (e.g. age, gender, disease characteristics, country, setting)	Assessed if all items are relevant for the study population in adequate sample size (≥ 10)	Assessed if all items are relevant for the study population in moderate sample size (5-9)	Assessed if all items are relevant for the study population in small sample size (<5)	NOT assessed if all items are relevant for the study population OR target population not involved
3	Was there an assessment of whether all items are relevant for the purpose of the measurement instrument? (discriminative, evaluative, and/or predictive)	Assessed if all items are relevant for the purpose of the application	Purpose of the instrument was not described but assumed	NOT assessed if all items are relevant for the purpose of the application	
4	Was there an assessment of whether all items together comprehensively reflect the construct to be measured?	Assessed if all items together comprehensively reflect the construct to be measured		No theoretical foundation of the construct and this was not taken into consideration	NOT assessed if all items together comprehensively reflect the construct to be measured
5	Were there any important flaws in the design or methods of the study?	No other important methodological flaws in the design or execution of the study		Other minor methodological flaws in the design or execution of the study	Other important methodological flaws in the design or execution of the study

Box E. Structural validity				
	excellent	good	fair	poor
1 Does the scale consist of effect indicators, i.e. is it based on a reflective model?				
<i>Design requirements</i>				
2 Was the percentage of missing items given?	Percentage of missing items described	Percentage of missing items NOT described		
3 Was there a description of how missing items were handled?	Described how missing items were handled	Not described but it can be deduced how missing items were handled	Not clear how missing items were handled	
4 Was the sample size included in the analysis adequate?	7* #items and ≥ 100	5* #items and ≥ 100 OR 5-7* #items but < 100	5* #items but < 100	$< 5^*$ #items
5 Were there any important flaws in the design or methods of the study?	No other important methodological flaws in the design or execution of the study		Other minor methodological flaws in the design or execution of the study (e.g. rotation method not described)	Other important methodological flaws in the design or execution of the study (e.g. inappropriate rotation method)

<i>Statistical methods</i>			
6	for CTT: Was exploratory or confirmatory factor analysis performed?	Exploratory or confirmatory factor analysis performed and type of factor analysis appropriate in view of existing information	Exploratory factor analysis performed while confirmatory would have been more appropriate
7	for IRT: Were IRT tests for determining the (uni-) dimensionality of the items performed?	IRT test for determining (uni)dimensionality performed	No exploratory or confirmatory factor analysis performed IRT test for determining (uni)dimensionality NOT performed

Box F. Hypotheses testing					
		excellent	good	fair	Poor
Design requirements					
1	Was the percentage of missing items given?	Percentage of missing items described	Percentage of missing items NOT described		
2	Was there a description of how missing items were handled?	Described how missing items were handled	Not described but it can be deduced how missing items were handled	Not clear how missing items were handled	
3	Was the sample size included in the analysis adequate?	Adequate sample size (≥100 per analysis)	Good sample size (50-99 per analysis)	Moderate sample size (30-49 per analysis)	Small sample size (<30 per analysis)

4	Were hypotheses regarding correlations or mean differences formulated a priori (i.e. before data collection)?	Multiple hypotheses formulated a priori	Minimal number of hypotheses formulate a priori	Hypotheses vague or not formulated but possible to deduce what was expected	Unclear what was expected
5	Was the expected <i>direction</i> of correlations or mean differences included in the hypotheses?	Expected direction of the correlations or differences stated	Expected direction of the correlations or differences NOT stated		
6	Was the expected absolute or relative <i>magnitude</i> of correlations or mean differences included in the hypotheses?	Expected magnitude of the correlations or differences stated	Expected magnitude of the correlations or differences NOT stated		
7	for convergent validity: Was an adequate description provided of the comparator instrument(s)?	Adequate description of the constructs measured by the comparator instrument(s)	Adequate description of most of the constructs measured by the comparator instrument(s)	Poor description of the constructs measured by the comparator instrument(s)	NO description of the constructs measured by the comparator instrument(s)
8	for convergent validity: Were the measurement properties of the comparator instrument(s) adequately described?	Adequate measurement properties of the comparator instrument(s) in a population similar to the study population	Adequate measurement properties of the comparator instrument(s) but not sure if these apply to the study population	Some information on measurement properties (or a reference to a study on measurement properties) of the comparator instrument(s) in any study population	No information on the measurement properties of the comparator instrument(s)

9	Were there any important flaws in the design or methods of the study?	No other important methodological flaws in the design or execution of the study	Other minor methodological flaws in the design or execution of the study (e.g. only data presented on a comparison with an instrument that measures another construct)	Other important methodological flaws in the design or execution of the study	
Statistical methods					
10	Were design and statistical methods adequate for the hypotheses to be tested?	Statistical methods applied appropriate	Assumable that statistical methods were appropriate, e.g. Pearson correlations applied, but distribution of scores or mean (SD) not presented	Statistical methods applied NOT optimal	Statistical methods applied NOT appropriate

Box G. Cross-cultural validity					
		excellent	good	fair	poor
Design requirements					
1	Was the percentage of missing items given?	Percentage of missing items described	Percentage of missing items NOT described		
2	Was there a description of how missing items were handled?	Described how missing items were handled	Not described but it can be deduced how missing items were handled	Not clear how missing items were handled	

3	Was the sample size included in the analysis adequate?	CTT: 7* #items and ≥100 IRT: ≥200 per group	CTT: 5* #items and ≥100 OR 5-7* #items but <100 IRT: ≥200 in 1 group and 100-199 in 1 group	CTT: 5* #items but <100 IRT: 100-199 per group	CTT: <5* #items IRT: (<100 in 1 or both groups
4	Were both the original language in which the HR-PRO instrument was developed, and the language in which the HR-PRO instrument was translated described?	Both source language and target language described			Source language NOT known
5	Was the expertise of the people involved in the translation process adequately described? e.g. expertise in the disease(s) involved, expertise in the construct to be measured, expertise in both languages	Expertise of the translators described with respect to disease, construct, and language	Expertise of the translators with respect to disease or construct poor or not described	Expertise of the translators with respect to language not described	
6	Did the translators work independently from each other?	Translators worked independent	Assumable that the translators worked independent	Unclear whether translators worked independent	Translators worked NOT independent
7	Were items translated forward and backward?	Multiple forward and multiple backward translations	Multiple forward translations but one backward translation	One forward and one backward translation	Only a forward translation
8	Was there an adequate description of how differences between the original and translated versions were resolved?	Adequate description of how differences between translators were resolved	Poorly or NOT described how differences between translators were resolved		

9	Was the translation reviewed by a committee (e.g. original developers)?	Translation reviewed by a committee (involving other people than the translators, e.g. the original developers)	Translation NOT reviewed by (such) a committee		
10	Was the HR-PRO instrument pre-tested (e.g. cognitive interviews) to check interpretation, cultural relevance of the translation, and ease of comprehension?	Translated instrument pre-tested in the target population	Translated instrument pre-tested, but unclear if this was done in the target population	Translated instrument pre-tested, but NOT in the target population	Translated instrument NOT pre-tested
11	Was the sample used in the pre-test adequately described?	Sample used in the pre-test adequately described		Sample used in the pre-test NOT (adequately) described	
12	Were the samples similar for all characteristics except language and/or cultural background?	Shown that samples were similar for all characteristics except language /culture	Stated (but not shown) that samples were similar for all characteristics except language /culture	Unclear whether samples were similar for all characteristics except language /culture	Samples were NOT similar for all characteristics except language /culture
13	Were there any important flaws in the design or methods of the study?	No other important methodological flaws in the design or execution of the study		Other minor methodological flaws in the design or execution of the study	Other important methodological flaws in the design or execution of the study

Statistical methods		
14 for CTT: Was confirmatory factor analysis performed?	Multiple-group confirmatory factor analysis performed	Multiple-group confirmatory factor analysis NOT performed
15 for IRT: Was differential item function (DIF) between language groups assessed?	DIF between language groups assessed	DIF between language groups NOT assessed

Box H. Criterion validity					
		excellent	good	fair	poor
Design requirements					
1	Was the percentage of missing items given?	Percentage of missing items described	Percentage of missing items NOT described		
2	Was there a description of how missing items were handled?	Described how missing items were handled	Not described but it can be deduced how missing items were handled	Not clear how missing items were handled	
3	Was the sample size included in the analysis adequate?	Adequate sample size (≥100)	Good sample size (50-99)	Moderate sample size (30-49)	Small sample size (<30)
4	Can the criterion used or employed be considered as a reasonable 'gold standard'?	Criterion used can be considered an adequate 'gold standard' (evidence provided)	No evidence provided, but assumable that the criterion used can be considered an adequate 'gold standard'	Unclear whether the criterion used can be considered an adequate 'gold standard'	Criterion used can NOT be considered an adequate 'gold standard'

5	Were there any important flaws in the design or methods of the study?	No other important methodological flaws in the design or execution of the study	Other minor methodological flaws in the design or execution of the study	Other important methodological flaws in the design or execution of the study
<i>Statistical methods</i>				
6	for continuous scores: Were correlations, or the area under the receiver operating curve calculated?	Correlations or AUC calculated		Correlations or AUC NOT calculated
7	for dichotomous scores: Were sensitivity and specificity determined?	Sensitivity and specificity calculated		Sensitivity and specificity NOT calculated

Box I. Responsiveness					
		excellent	good	fair	poor
<i>Design requirements</i>					
1	Was the percentage of missing items given?	Percentage of missing items described	Percentage of missing items NOT described		
2	Was there a description of how missing items were handled?	Described how missing items were handled	Not described but it can be deduced how missing items were handled	Not clear how missing items were handled	
3	Was the sample size included in the analysis adequate?	Adequate sample size (≥100)	Good sample size (50-99)	Moderate sample size (30-49)	Small sample size (<30)
4	Was a longitudinal design with at least two measurement used?	Longitudinal design used			No longitudinal design used
5	Was the time interval stated?	Time interval adequately described			Time interval NOT described

6	If anything occurred in the interim period (e.g. intervention, other relevant events), was it adequately described?	Anything that occurred during the interim period (e.g. treatment) adequately described	Assumable what occurred during the interim period	Unclear or NOT described what occurred during the interim period
7	Was a proportion of the patients changed (i.e. improvement or deterioration)?	Part of the patients were changed (evidence provided)	NO evidence provided, but assumable that part of the patients were changed	Unclear if part of the patients were changed Patients were NOT changed
Design requirements for hypotheses testing				
For constructs for which a gold standard was not available:				
8	Were hypotheses about changes in scores formulated a priori (i.e. before data collection)?	Hypotheses formulated a priori		Hypotheses vague or not formulated but possible to deduce what was expected Unclear what was expected
9	Was the expected <i>direction</i> of correlations or mean differences of the change scores of HR-PRO instruments included in these hypotheses?	Expected direction of the correlations or differences stated	Expected direction of the correlations or differences NOT stated	
10	Were the expected absolute or relative <i>magnitude</i> of correlations or mean differences of the change scores of HR-PRO instruments included in these hypotheses?	Expected magnitude of the correlations or differences stated	Expected magnitude of the correlations or differences NOT stated	

11	Was an adequate description provided of the comparator instrument(s)?	Adequate description of the constructs measured by the comparator instrument(s)	Poor description of the constructs measured by the comparator instrument(s)	NO description of the constructs measured by the comparator instrument(s)
12	Were the measurement properties of the comparator instrument(s) adequately described?	Adequate measurement properties of the comparator instrument(s) in a population similar to the study population	Adequate measurement properties of the comparator instrument(s) but not sure if these apply to the study population	Some information on measurement properties (or a reference to a study on measurement properties) of the comparator instrument(s) in any study population
13	Were there any important flaws in the design or methods of the study?	No other important methodological flaws in the design or execution of the study	Other minor methodological flaws in the design or execution of the study (e.g. only data presented on a comparison with an instrument that measures another construct)	Other important methodological flaws in the design or execution of the study
<i>Statistical methods</i>				
14	Were design and statistical methods adequate for the hypotheses to be tested?	Statistical methods applied appropriate	Statistical methods applied NOT optimal	Statistical methods applied NOT appropriate

Design requirement for comparison to a gold standard					
For constructs for which a gold standard was available:					
15	Can the criterion for change be considered as a reasonable gold standard?	Criterion used can be considered an adequate 'gold standard' (evidence provided)	No evidence provided, but assumable that the criterion used can be considered an adequate 'gold standard'	Unclear whether the criterion used can be considered an adequate 'gold standard'	Criterion used can NOT be considered an adequate 'gold standard'
16	Were there any important flaws in the design or methods of the study?	No other important methodological flaws in the design or execution of the study		Other minor methodological flaws in the design or execution of the study	Other important methodological flaws in the design or execution of the study
Statistical methods					
17	for continuous scores: Were correlations between change scores, or the area under the Receiver Operator Curve (ROC) curve calculated?	Correlations or Area under the ROC Curve (AUC) calculated			Correlations or AUC NOT calculated
18	for dichotomous scales: Were sensitivity and specificity (changed versus not changed) determined?	Sensitivity and specificity calculated			Sensitivity and specificity NOT calculated

Interpretability

We recommend to use the Interpretability box to extract all information on the interpretability issues described in this box of the instruments under study from the included articles.

Box Interpretability	
Percentage of missing items	
Description of how missing items were handled	
Distribution of the (total) scores	
Percentage of the respondents who had the lowest possible (total) score	
Percentage of the respondents who had the highest possible (total) score	
Scores and change scores (i.e. means and SD) for relevant (sub) groups, e.g. for normative groups, subgroups of patients, or the general population	
Minimal Important Change (MIC) or Minimal Important Difference (MID)	

Generalizability

We recommend to use the Generalizability box to extract data on the characteristics of the study populations and sampling procedures of the included studies.

Box Generalisability	
Median or mean age (with standard deviation or range)	
Distribution of sex	
Important disease characteristics (e.g. severity, status, duration) and description of treatment	
Setting(s) in which the study was conducted (e.g. general population, primary care or hospital/rehabilitation care)	
Countries in which the study was conducted	
Language in which the HR-PRO instrument was evaluated	
Method used to select patients (e.g. convenience, consecutive, or random)	
Percentage of missing responses (response rate)	

Appendix 2: Search-Construct SVMC measurement-tool list

Pubmed

(selective voluntary motor control[TIAB] OR selective movement*[TIAB] OR Selectivity[TIAB] OR voluntary movement*[TIAB] OR volitional movement*[TIAB] OR quality of movement*[TIAB] OR muscle activation[TIAB] OR muscular activation[TIAB] OR movement pattern*[TIAB] OR physiological movement*[TIAB] OR sensorimotor control[TW] OR foot targeting[TW] OR foot placing[TW] OR coordinated movement*[TIAB] OR motion compensation OR compensatory movement* OR muscular coactivation OR mirror movement* OR muscle synerg* OR muscular synerg* OR movement synerg* OR cortical overlap[TW] OR motor control[TW] OR joint coupling[TW] OR (interjoint OR inter-joint AND coordination) OR force modulation[TW] OR submaximal contraction*[TW] OR maximal voluntary contraction*[TW]) OR ("Muscle, Skeletal"[MeSH] OR "Motor Skills Disorders"[Mesh] OR "Neuromuscular Manifestations"[Mesh] OR "Motor Activity"[Mesh] OR "Recovery of Function/physiology"[Mesh] OR "Gait Disorders, Neurologic"[Mesh])) AND (upper motor neuron lesion*[TW] OR "Cerebral Palsy"[Mesh] OR "Brain Injuries"[Mesh] OR "Stroke"[Mesh] OR "spinal cord injuries"[MeSH])) AND ((instrumentation[sh] OR methods[sh] OR Validation Studies[pt] OR Comparative Study[pt] OR "psychometrics"[MeSH] OR psychometr*[tiab] OR clinimetr*[tw] OR clinometr*[tw] OR "outcome assessment (health care)"[MeSH] OR outcome assessment[tiab] OR outcome measure*[tw] OR "observer variation"[MeSH] OR observer variation[tiab] OR "Health Status Indicators"[Mesh] OR "reproducibility of results"[MeSH] OR reproducib*[tiab] OR "discriminant analysis"[MeSH] OR reliab*[tiab] OR unreliab*[tiab] OR valid*[tiab] OR coefficient[tiab] OR homogeneity[tiab] OR homogeneous[tiab] OR "internal consistency"[tiab] OR (cronbach*[tiab] AND (alpha[tiab] OR alphas[tiab])) OR (item[tiab] AND (correlation*[tiab] OR selection*[tiab] OR reduction*[tiab])) OR agreement[tiab] OR precision[tiab] OR imprecision[tiab] OR "precise values"[tiab] OR test-retest[tiab] OR (test[tiab] AND retest[tiab]) OR (reliab*[tiab] AND (test[tiab] OR retest[tiab])) OR stability[tiab] OR interrater[tiab] OR inter-rater[tiab] OR intrarater[tiab] OR intra-rater[tiab] OR intertester[tiab] OR inter-tester[tiab] OR intratester[tiab] OR intra-tester[tiab] OR interobserver[tiab] OR inter-observer[tiab] OR intraobserver[tiab] OR intra-observer[tiab] OR intertechnician[tiab] OR inter-technician[tiab] OR intratechnician[tiab] OR intra-technician[tiab] OR interexaminer[tiab] OR inter-examiner[tiab] OR intraexaminer[tiab] OR intra-examiner[tiab] OR interassay[tiab] OR inter-assay[tiab] OR intraassay[tiab] OR intra-assay[tiab] OR interindividual[tiab] OR inter-individual[tiab] OR intraindividual[tiab] OR intra-individual[tiab] OR interparticipant[tiab] OR inter-participant[tiab] OR intraparticipant[tiab] OR intra-participant[tiab] OR kappa[tiab] OR kappa's[tiab] OR kappas[tiab] OR repeatab*[tiab] OR ((replicab*[tiab] OR repeated[tiab]) AND (measure[tiab] OR measures[tiab] OR findings[tiab] OR result[tiab] OR results[tiab] OR test[tiab] OR tests[tiab])) OR generaliza*[tiab] OR generalisa*[tiab] OR concordance[tiab] OR (intraclass[tiab] AND correlation*[tiab]) OR discriminative[tiab] OR "known

group"[tiab] OR factor analysis[tiab] OR factor analyses[tiab] OR dimension*[tiab] OR subscale*[tiab] OR (multitrait[tiab] AND scaling[tiab] AND (analysis[tiab] OR analyses[tiab])) OR item discriminant[tiab] OR interscale correlation*[tiab] OR error[tiab] OR errors[tiab] OR "individual variability"[tiab] OR (variability[tiab] AND (analysis[tiab] OR values[tiab])) OR (uncertainty[tiab] AND (measurement[tiab] OR measuring[tiab])) OR "standard error of measurement"[tiab] OR sensitiv*[tiab] OR responsive*[tiab] OR ((minimal[tiab] OR minimally[tiab] OR clinical[tiab] OR clinically[tiab]) AND (important[tiab] OR significant[tiab] OR detectable[tiab]) AND (change[tiab] OR difference[tiab])) OR (small*[tiab] AND (real[tiab] OR detectable[tiab]) AND (change[tiab] OR difference[tiab])) OR meaningful change[tiab] OR "ceiling effect"[tiab] OR "floor effect"[tiab] OR "Item response model"[tiab] OR IRT[tiab] OR Rasch[tiab] OR "Differential item functioning"[tiab] OR DIF[tiab] OR "computer adaptive testing"[tiab] OR "item bank"[tiab] OR "cross-cultural equivalence"[tiab])))) AND Humans[Mesh]) NOT ("Randomized Controlled Trial"[Publication Type] OR "Controlled Clinical Trial"[Publication Type] OR "addresses"[Publication Type] OR "biography"[Publication Type] OR "case reports"[Publication Type] OR "comment"[Publication Type] OR "directory"[Publication Type] OR "editorial"[Publication Type] OR "festschrift"[Publication Type] OR "interview"[Publication Type] OR "lectures"[Publication Type] OR "legal cases"[Publication Type] OR "legislation"[Publication Type] OR "letter"[Publication Type] OR "news"[Publication Type] OR "newspaper article"[Publication Type] OR "patient education handout"[Publication Type] OR "popular works"[Publication Type] OR "congresses"[Publication Type] OR "consensus development conference"[Publication Type] OR "consensus development conference, nih"[Publication Type] OR "practice guideline"[Publication Type] OR "intervention studies"[MESH])) AND Humans [Mesh] AND child [Mesh]

Appendix 3: Overview of SVMC measurement-tools

Please see next page for the results of the initial search (please see Appendix 2 for the search term) aiming to find available SVMC measurement-tools for patients with UMN lesions.

Appendix 3: Overview of SVMC measure for the lower extremity in patients with UMN lesions

SVMC Outcome	Population					Study design		Body part	SVMC scoring
	CP	TBI	CVI	SCI	H	Psycho.	Other		
observation based SVMC measures									
Therapist observation [Mayo et al. 1991] video-taped observations [Pomeroy et al. 2003]		X	X			R		LE	neurophysiological appearance of voluntary movements
ASIA Lower Extremity motor score (LMES) [Graves et al. 2006; Savic et al. 2007; Furlan et al. 2011; Hadley et al. 2013; Villiger et al. 2013]				X		R,V	Interv.	LE and toes	lower motor score (LEMS) (ordinal 5-0)
Fugl Meyer Assessment (FMA) [Fugl-Meyer et al. 1975; 1980; Duncan et al. 1994; Van der Lee et al. 2001; Gladstone et al. 2002; 2015; Hsieh et al. 2007; Woodbury et al. 2007; Crow et al. 2008; Bowden et al. 2010; Knutson and Chae 2011; Kiyama et al. 2011 Mirbagheri et al. 2011; Hiengkaew et al. 2012; Tang et al. 2012; Chen et al. 2014]			X			R,V	Interv.	knee, ankle	items III, IV ordinal (2-0) also test in combi. with EMG
Brunnstrom recovery stages [Chen et al. 2003; Sütbeyaz et al. 2007]			X				Cohort Interv.	LE	subscale LE 6 stages
Stroke Rehabilitation Assessment of Movement ((S-)STREAM) [Daley et al. 1999; Wang et al. 2002; Ahemd et al. 2003; Huesch et al. 2006; Hsieh et al. 2008; Waldmann et al. 2013]			X			R,V, RES	Cohort	LE	subscale 2: voluntary motor ability of the LE ordinal (2-0) MCID value 1.9; SRM=1.14
Selective Motor Control Test (SMC) [Ostensjo et al. 2007; Lowing and Carlberg; 2009; Hayek et al. 2009; Smits et al. 2010; Park and Kim 2013]	X					R	Cohort, Interv.	ankle	ordinal (2-0)
Modified Trost Selective Motor Control Test (mTrost) [Desloovere 2006; Voorman et al. 2007; Smits et al. 2010]	X					R	Cohort	LE	ordinal (4-0)
Selective Control Assessment of the lower extremity (SCALE) [Fowler et al. 2009; [Fowler and Goldberg 2009; Goldberg et al. 2010; Goldberg et al. 2011, 2012; Wu et al. 2011; Sukal-Moulton et al. 2014]	X					R,V	Cohort, Interv.	LE and STJ, toes	ordinal (2-0)
laboratory based SVMC measures									
Foot tapping test (FTT) [Kawahira et al. 2004; Numasawa et al. 2012]									
Rhythm Task (video & metronome) [Tomita and Usuda 2013]			X	X	X		Cohort Interv.	ankle	number of foot taps in 30 sec.
kinematic measures									
Active range of motion [Toner et al. 1998; Knutson and Chae 2011]	X						Interv.	ankle	degrees of dorsiflexion (biofeedback training)
Electro-goniometer [Byaton et al. 2006; Beaulieu et al. 2014]	X						Interv.	ankle	virtual exercise program
Video kinematic sagittal [Olree et al. 2000]	X				X		Interv.	leg	correlation synergy-movement
Passive marker system [Engsberg et al. 2004, 2008]	X				X REF	R	Cohort	ankle	SD of relative ankle phase; during 4 tasks
Robotic Systems [Villger et al. 2013; Cioi et al. 2011; Wu et al. 2011; Burdea et al. 2013]					X		Cohort Interv.	knee, ankle	Correlation coefficient ideal path / actual movement
force measure									
isometric muscle force / torque (plus EMG) [Lebiedowska and Fisk 2003; Lebiedowska et al. 2004; Buckon et al. 2002]	X				X		Interv.	knee, ankle	maximal voluntary contractions and co-contraction post SDR
Dynamometer [Engsberg et al. 2000; Damiano et al. 2000; Beaulieu et al. 2014]	X				X REF		Cohort	knee, ankle	target force: mean, SD, median frequency ; co-contraction ratio
Temporal force/torque characteristics [Downing et al. 2009; Varoqui et al. 2014]	X			X	X		Interv. Cohort	LE	MVC, reaction time, time of movement production,
EMG measures									
On-Off ratio, poly-electromyography [Lavoie et al. 1997; Berger 1998; Rose et al. 1999; Chen et al. 2000; Damiano et al. 2000; Policy et al. 2001; Maegele et al. 2002; Calancie et al. 2004; Romkes & Bruner 2007; Olsen et al. 2013; Tang et al. 2012] muscle activation latencies/ ratios /cross correlations [Daly et al. 2011; Tedroff et al. 2008; van Doornik et al. 2008; Zwaan et al. 2012; Rios et al. 2013]	X			X	X		Cohort Interv.	muscle LE	co-contraction (maturation walking); voluntary tasks, neuro-gametherapy
Frequency analysis walking [Lauer et al. 2005; 2007a,b; Prosser et al. 2010] root mean square; integrated EMG [Wang et al. 2011]	X						Cohort Interv.	muscle LE	walking, voluntary tasks post-op
Amplitude [Patikas et al. 2007]	X						Interv.	muscle LE	walking, post-multilevel op
Voluntary response index (VRI) [Lim et al. 2004; Lee et al. 2004; McKay et al.2004; 2005; Zoghi et al. 2013]				X		R	Cohort	muscle LE	Voluntary response index during voluntary movement
Submaximal force tracking									
Visumotor tracking [Jayaraman et al. 2006; Wirth et al. 2008; van Hedel et al. 2011; Knutson and Chae 2011; Kiyama et al. 2011]			X	X		R	Cohort Interv.	knee, ankle	root mean square (RMS) tracking performance; smoothness of movement
Torque steadiness [Bandholm et al. 2008; 2009; 2011; Arpin et al.2013]	X				X	R	Cohort Interv.	muscle LE	torque SD; variability, Coefficient of variance
Transcranial Magnet Stimulation (TMS)									
Motor Evoked Potentials (MEPs) [Brouwer & Ashby 1991; Maegele et al. 2002; Jensen et al. 2005; Wirth et al. 2008a,b,c,d; Garvey & Mall 2008; Dimyan et al. 2010; Beaulieu et al. 2014; Lewis et al. 2014]	X	X	X		X	R	Cohort	Ankle (TA)	magnitude; latencies MEP; SICI;SICF; during active DE and walking
Motor Mapping [Wittenberg et al. 2009; Ziemann 2011; Kesar et al. 2014]	X				X		Cohort	Ankle (TA)	center of gravity (COG) and the number of active positions (TA)
Diffusion Tensor Imaging (DTI) & functional Magnet Resonance Imagining (fMRI)									
[Staudt et al. 2003; Dobkin et al. 2004;Holodny et al. 2005; Thomas et al. 2005; Christensen et al. 2007; Ludeman et al. 2008; Jurkiewicz et al. 2010; Hoon et al. 2010; Koerte et al. 2011; Lee et al. 2011]	X		X		X REF		Cohort	Ankle	fractional anisotropy and trace; volume, voxel correlation
Abbreviations: CP: Cerebral Palsy; TBI: Traumatic Brain Injury; SCI: Spinal Cord Injury; H: Healthy controls; Psycho.: Psychometric Property assessed; R: Reliability; V: Validity; RES: Responsiveness; Interv.: Intervention; LE: Lower Extremity joints; hip, knee, ankle; TA: m. Tibialis Anterior									

Appendix 4: Study - German Version SCALE

Please see next pages.

SCALE: Selective Control Assessment of the Lower Extremity
„Untersuchung der selektiven Kontrolle der unteren Extremität“
SCALE Auswertungsformular

Datum: _____ ID: _____ Alter: _____ GMFCS Level: _____

Diagnose: ☐ Spastische Diplegie ☐ Spastische Tetraplegie ☐ Spastische Hemiplegie R L Andere: _____

Links						Rechts				
Bewertung	Hüfte	Knie	OSG	USG	Zeh- en	Hüfte	Knie	OSG	USG	Zeh- en
Normal (2 Punkte)										
Beeinträchtigt (1 Punkt)										
Unfähig (0 Punkte)										
Total Bein Score	L =	R =								

Synergie ausgelöst durch Widerstand

Knieextension mit Widerstand gegen Beinextensionsmuster										
Dorsalflexion mit Widerstand gegen Beinflexionsmuster										

Merkmale

Hüftflexionskontraktur										
Adduktorenkontraktur oder Spastizität										
Knieflexionskontraktur										
Verkürzung der ischiocruralen Muskulatur										
Plantarflexionskontraktur										
Plantarflexoren Spastizität										
Inversion oder Eversion, keine reine Dorsalflexion										
Bewegt primär die Zehen										
Spiegelbewegungen am kontralateralen Bein										
Bewegt langsamer als im Dreisekudentakt										
Bewegt nur in eine Richtung (notiere die erreichte Bewegung)										
Mitbewegung von anderen Gelenken										
Bewegung ≤ 50% des verfügbaren (passiven) ROM										

Kommentare zum Test: _____

Die deutsche Version der SCALE basiert auf der englischen Originalversion (Fowler et al. Dev. Med. Child. Neurol. 2009;51:607-14) und wurde im Rahmen der Doktorarbeit von Frau Julia Balzer erstellt
© Copyright 2009 englische Originalversion 2.3 - University of California Los Angeles/ Orthopedic Hospital Center for Cerebral Palsy
© Copyright 2013 für die deutsche Version – Universitäts-Kinderspital Zürich – Eleonorenstiftung

Untersucher

SCALE: Anleitung zur Ausführung

Allgemeine Bemerkung: Zugunsten der Lesbarkeit sind alle personenbezogenen Bezeichnungen in männlicher Form gehalten, gelten aber ausdrücklich für beide Geschlechter. Patienten unter 16 Jahren können in einer persönlicheren Form (Du) angesprochen werden. Begriffe, welche im Folgenden in Klammern (...) aufgeführt sind, sind nicht in der Originalversion enthalten, sondern wurden für ein besseres Verständnis zusätzlich eingefügt.

Der Patient muss fähig sein, einfache motorische Anweisungen zu folgen. Um diese Fähigkeit zu testen, bitten Sie den Patienten den am wenigsten betroffenen Körperteil zu bewegen. Bevor Sie den Patient bitten, das jeweilige Gelenk zu bewegen, soll jedes Gelenk passiv bewegt werden, um das gesamte Bewegungsausmass zu erfassen. Um sicher zu gehen, dass der Patient die Aufgaben richtig versteht, demonstrieren Sie die Bewegungssequenz während Sie gleichzeitig das Bein unterstützen. Die Instruktionen an den Patienten, sind Vorschläge und dürfen angepasst werden, um eine optimale Ausführung bei jedem individuellen Patienten hervorzurufen. Um den Patienten in der gewünschten Bewegungs-Geschwindigkeit anzuleiten, soll während der Aufgabe im Dreisekudentakt gezählt werden. Mehrere Versuche sind gestattet, des Weiteren sind Rückmeldungen, um die Ausführung zu verbessern, zugelassen.

Allgemeine Instruktionen an den Patienten - „Ich werde Sie bitten, sich auf eine bestimmte Weise zu bewegen. Versuchen Sie, sich so zu bewegen, wie ich Sie anleite. Versuchen Sie keinen anderen Körperteil zu bewegen. Wenn Sie Fragen haben oder unsicher sind, ob sie mich richtig verstanden haben, sagen Sie mir bitte Bescheid.“

Hüfte

Ausgangsstellung: In der Seitenlage, Hüfte und Knie sind in voller Extension. Unterstützen Sie das (obere, zu testende) Bein medial am Knie- und Fussgelenk. Zu Stabilitätszwecken darf das unten liegende, nicht zu testende Bein, flektiert werden. Die zu testende Bewegung ist die Hüftflexion, wobei das Knie gestreckt bleiben soll. Erfassen Sie das (passive) Bewegungsausmass der Hüftflexion mit gestrecktem Knie, da es durch eine Verkürzung der ischiocruralen Muskulatur eingeschränkt sein könnte.

Wenn der Patient durch eine Verkürzung der ischiocruralen Muskulatur Schwierigkeiten mit dieser Aufgabe hat, dann bitten Sie ihn, die Hüfte zu strecken, zu beugen und wieder zu strecken während das Knie (des zu testenden / oberen Beines) in 90° Flexion gehalten wird. Um die Ausführung dieser Aufgabe gut evaluieren zu können, ist eine adäquate Erfassung des gesamten (passiven) Bewegungsausmasses der Hüftextension notwendig.

Instruktionen an den Patienten – Bitten Sie den Patienten die Hüfte zu beugen, zu strecken und (wieder) zu beugen, während das Knie gestreckt bleibt. Zum Beispiel: „Bewegen Sie ihr Bein nach vorne, zurück und wieder nach vorne, während Sie ihr Knie gestreckt halten. Ich werde Sie erst durch die Bewegung führen, und dann möchte ich Sie bitten dies selber zu versuchen.“

Knie

Ausgangsstellung: Die übrigen Tests werden im Sitzen durchgeführt, wobei die Unterschenkel über den Rand der Untersuchungsliege hängen. Während der folgenden Übungen, darf der Patient sich, ungefähr 20° zurücklehnen und sich auf den Händen abstützen, um für eine Verkürzung der ischiocruralen Muskulatur zu kompensieren.

Instruktionen an den Patienten – Bitten Sie den Patienten das Knie zu strecken, zu beugen und dann wieder zu strecken während die Hüfte gebeugt bleibt. Zum Beispiel: „Strecken Sie ihr Knie so gut wie möglich, dann beugen Sie es und strecken es anschliessend wieder. Versuchen Sie dies zu tun ohne weiter nach hinten zu lehnen oder Ihr anderes Bein mitzubewegen. Ich werde Sie erst durch die Bewegung führen, und dann möchte ich Sie bitten dies selber zu versuchen.“

Extensionssynergie des Beines: Falls eine Quadrizepsschwäche vermutet wird, kann die Extensionssynergie des Beines getestet werden. Erlauben Sie dem Patienten sich zurück zu lehnen und sich auf seine Hände abzustützen oder sich anzulehnen, sodass der Rumpf etwa 45° zurückgelehnt ist. Positionieren Sie das Bein in Hüft- und Knieflexion mit dem Fussgelenk in Dorsalextension. Bitten Sie den Patienten gegen Ihre Hand zu drücken, das Knie zu strecken und Fussgelenk und Zehen zu plantarflektieren. Geben Sie Widerstand an den Metatarsalköpfchen und

vergleichen Sie das Ausmass der Kniestreckung mit dem Ausmass, welches beim (vorherigen) selektiven Willkürmotoriktest des Knies erreicht worden ist.

Fussgelenk (OSG)

Ausgangsstellung: Im Sitzen, wie im (vorherigen) Knieextensions-Test. Das Knie ist gestreckt und der Untersucher unterstützt den Unterschenkel. Erfassen Sie das passive Dorsalextension-Bewegungsausmass des Fussgelenkes während das Kniegelenk gestreckt ist. Das Knie darf bis zu 20° gebeugt werden, um für eine Verkürzung der ischiocruralen Muskulatur und/oder des Gastrocnemius zu kompensieren.

Instruktionen an den Patienten: Bitten Sie den Patienten das Fussgelenk zu dorsalflektieren, plantarflektieren und dann (wieder) zu dorsalflektieren während das Knie in gestreckt bleibt. Zum Beispiel: „Während ich Ihr Bein unterstütze und Sie Ihr Knie gestreckt halten, versuchen Sie Ihren Fuss nach oben, nach unten und wieder nach oben zu bewegen. Ich werde Sie erst durch die Bewegung führen, und dann möchte ich Sie bitten dies selber zu versuchen.“

Flexionssynergietest des Beines „Konfusionstest“¹: Wenn eine Schwäche der Dorsalextension vermutet wird, kann die Flexionssynergie des Beines untersucht werden. Bitten Sie den Patienten die Hüfte zu beugen, während das Knie gebeugt bleiben soll. Geben Sie Widerstand gegen Hüftflexion distal am Oberschenkel. Vergleichen Sie das Bewegungsausmass der Dorsalflexion mit dem Bewegungsausmass, welches beim Testen der selektiven Willkürbewegung des Fussgelenkes erreicht worden ist.

Fuss / Unteres Sprunggelenk (USG)

Ausgangsstellung: Im Sitzen, wie in den (vorherigen) Knie- und Fussgelenktests. Die Wade wird unterstützt gehalten.

Instruktionen an den Patienten – Bitten Sie den Patienten den Fuss zu invertieren, evertieren, dann wieder zu invertieren während das Knie gestreckt bleibt. Zum Beispiel: „Bewegen Sie Ihren Fuss nach innen, aussen und dann wieder nach innen, während ich Ihr Bein unterstütze. Ich werde Sie erst durch die Bewegung führen, und dann möchte ich Sie bitten dies selber zu versuchen.“

Zehen

Ausgangsstellung: Im Sitzen, wie im (vorherigen) Fussgelenktest. Die Ferse wird unterstützt gehalten.

Instruktionen an den Patienten – Bitten Sie den Patienten die Zehen zu beugen, zu strecken und wieder zu beugen ohne das Fussgelenk oder das Knie zu bewegen. Zum Beispiel: „Beugen Sie alle Ihre Zehen nach unten, dann strecken Sie sie wieder und beugen Sie dann wieder nach unten, während ich Ihr Bein unterstütze. Ich werde Sie erst durch die Bewegung führen, und dann möchte ich Sie bitten dies selber zu versuchen.“

¹ Davis et al. *J Pediatr Orthop.* 1993;13(6):717-21.

Die deutsche Version der SCALE basiert auf der englischen Originalversion (Fowler et al. *Dev. Med. Child. Neurol.* 2009;51:607-14) und wurde im Rahmen der Doktorarbeit von Frau Julia Balzer erstellt
 © Copyright 2009 englische Originalversion 2.3 - University of California Los Angeles/ Orthopedic Hospital Center for Cerebral Palsy
 © Copyright 2013 für die deutsche Version – Universitäts-Kinderspital Zürich – Eleonorenstiftung

SCALE: Selective Control Assessment of the Lower Extremity **„Untersuchung der selektiven Kontrolle der unteren Extremität“**

Anleitung für die Bewertung

Jedes Gelenk wird mit 2, 1 oder 0 Punkten bewertet. Diese Punkte werden summiert und ergeben die Total Score für jedes Bein. Die Anzahl der Punkte für jede Bewertungsstufe steht in Klammern. Notieren Sie für jedes Gelenk, die Gelenks-Punktzahl und alle verfügbaren beschreibenden Merkmale auf dem SCALE Auswertungsformular.

Hüfte

Normal (2)

Beugt, streckt und beugt wieder. (Diese Bewegungen sollen flüssig mit einer normalen Bewegungsgeschwindigkeit stattfinden. Die Bewegungen finden ungefähr in einem Sekundentakt statt; somit dauert der Test ca. 3 Sekunden). Die Hüftbeugung erfolgt ohne Kniebeugung, innerhalb der normalen Bewegungsgeschwindigkeit und ohne Spiegelbewegungen [die gleiche Bewegung erfolgt am anderen Bein]. Wenn alternativ der Hüftextensionstest durchgeführt wurde: streckt, beugt und streckt dann wieder. Die Hüftstreckung erfolgt ohne Kniestreckung, innerhalb der normalen Bewegungsgeschwindigkeit und ohne Spiegelbewegungen.

Beeinträchtigt (1)

Eine oder mehrere der folgende Dinge treten auf: streckt oder beugt $\leq 50\%$ des verfügbaren passiven Bewegungsausmasses in der Testposition, die Aufgabe wird langsamer als in drei Sekunden durchgeführt, Spiegelbewegungen treten auf, die Bewegung erfolgt nur in eine Richtung oder eine Bewegung tritt an einem nicht getesteten Gelenk auf.

Unfähig (0)

Die Hüfte wird nicht gebeugt oder gestreckt oder mit gleichzeitiger Mitbewegung des Knies.

Knie

Normal (2)

Streckt, beugt und streckt wieder. (Diese Bewegungen sollen flüssig mit einer normalen Bewegungsgeschwindigkeit stattfinden. Die Bewegungen finden ungefähr in einem Sekundentakt statt; somit dauert der Test ca. 3 Sekunden). Die Bewegung erfolgt innerhalb der normalen Bewegungsgeschwindigkeit, ohne Mitbewegung des Rumpfes oder anderer Gelenke und ohne Spiegelbewegungen. Die Bewertung „normal“ darf gegeben werden, wenn das Knie in der Testposition über $> 50\%$ des passiven Bewegungsausmasses bewegt wird.

Beeinträchtigt (1)

Eine oder mehrere der folgende Dinge treten auf: streckt $\leq 50\%$ des verfügbaren passiven Bewegungsausmasses in der Ausgangstellung, die Aufgabe wird langsamer als in drei Sekunden durchgeführt, Spiegelbewegungen treten auf, die Bewegung erfolgt nur in eine Richtung oder eine Bewegung tritt an einem nicht getesteten Gelenk auf.

Unfähig (0)

Das Knie wird nicht gestreckt oder es treten gleichzeitig Hüft- oder Fussgelenksbewegungen.

Fussgelenk (OSG)

Normal (2)

Dorsalflektiert, plantarflektiert und dorsalflektiert wieder. (Diese Bewegungen sollen flüssig mit einer normalen Bewegungsgeschwindigkeit stattfinden. Die Bewegungen finden ungefähr in einem Sekundentakt statt; somit dauert der Test ca. 3 Sekunden). Die Bewegung innerhalb der normalen Bewegungsgeschwindigkeit, ohne Mitbewegung anderer Gelenke und ohne Spiegelbewegungen. Es müssen mindestens 15° Bewegung im Fussgelenk in der Sagittalebene beobachtet werden.

Beeinträchtigt (1)

Eine oder mehrere der folgenden Dinge treten auf: dorsalflektiert $\leq 50\%$ des verfügbaren passiven Bewegungsausmasses in der Testposition oder vom aktiven Bewegungsausmass während der Flexionssynergie des Beines, die Aufgabe wird langsamer als in drei Sekunden durchgeführt, Spiegelbewegungen treten auf, die

Bewegung erfolgt nur in eine Richtung oder eine Bewegung tritt an einem nicht getesteten Gelenk auf. Die Bewertung „beeinträchtigt“ wird gegeben, wenn gleichzeitig zu der Bewegung (des Fussgelenkes / Dorsalflexion) eine Streckung der Zehen oder eine Inversion im Fussgelenk stattfindet.

Unfähig (0)

Es findet keine Dorsalflexion statt oder nur mit gleichzeitiger Hüft- und Kniebeugung.

Unteres Sprunggelenk (USG)

Normal (2)

Invertiert, evertiert und invertiert wieder. (Diese Bewegungen sollen flüssig mit einer normalen Bewegungsgeschwindigkeit stattfinden. Die Bewegungen finden ungefähr in einem Sekundentakt statt; somit dauert der Test ca. 3 Sekunden). Die Bewegung innerhalb der normalen Bewegungsgeschwindigkeit ohne Mitbewegung anderer Gelenke und ohne Spiegelbewegungen. Eine aktive Eversion muss stattfinden.

Beeinträchtigt (1)

Eine oder mehrere der folgende Dinge treten auf: invertiert oder evertiert ≤ 50 des verfügbaren passiven Bewegungsausmasses in der Testposition, die Aufgabe wird langsamer als in drei Sekunden durchgeführt, Spiegelbewegungen treten auf, die Bewegung erfolgt nur in eine Richtung oder eine Bewegung tritt an einem nicht getesteten Gelenk auf.

Unfähig (0)

Es findet keine Inversion oder Eversion statt oder die Bewegung tritt nur im Synergie-Muster auf. Es tritt eine Dorsalflexion und/oder Plantarflexion oder keine Bewegung auf.

Zehen

Normal (2)

Beugt, streckt und beugt wieder. (Diese Bewegungen sollen flüssig mit einer normalen Bewegungsgeschwindigkeit stattfinden. Die Bewegungen finden ungefähr in einem Sekundentakt statt; somit dauert der Test ca. 3 Sekunden). Die Bewegung innerhalb der normalen Bewegungsgeschwindigkeit, ohne Mitbewegung anderer Gelenke und ohne Spiegelbewegungen. Alle fünf Zehen müssen sich bewegen.

Beeinträchtigt (1)

Eine oder mehrere der folgende Dinge treten auf: beugt oder streckt $\leq 50\%$ des verfügbaren passiven Bewegungsausmasses in der Testposition, die Aufgabe wird langsamer als in drei Sekunden durchgeführt, Spiegelbewegungen treten auf, die Bewegung erfolgt nur in eine Richtung oder eine Bewegung tritt an einem nicht getesteten Gelenk auf.

Unfähig (0)

Zehen werden nicht gebeugt **und nicht** gestreckt.

Unterschied zwischen „Unfähig“ und „Beeinträchtigt“

Unfähig (totale Synergien) beinhaltet, dass sich zwei oder mehrere Gelenken gleichzeitig mitbewegen. Für jedes Grad der Bewegung am zu untersuchenden Gelenk, treten zwingend gleichzeitig Mitbewegungen in anderen Gelenken des gleichen Beines auf, welche Teil eines Syngiemusters sind. Patienten mit beeinträchtigter motorischer Kontrolle sind fähig das zu untersuchende Gelenk über ein kleines Bewegungsausmass zu bewegen, ohne dass sich andere Gelenke mitbewegen, aber während einem Teil der Bewegung wird die Bewegung von Mitbewegungen in einem benachbartem Gelenk begleitet.

Unterschied zwischen „Beeinträchtigt“ und „Normal“

Normal motorische Kontrolle ist die Fähigkeit, ein einzelnes Gelenk durch mehr als 50% des verfügbaren Bewegungsausmasses innerhalb der normalen Bewegungsgeschwindigkeit in einer alternierenden Weise isoliert zu bewegen (Dreisekundentakt). Die Bewegung erfolgt ohne Mitbewegung von anderen Gelenken der Beine. Die Unfähigkeit, Bewegung in dieser Art auszuführen, wird als „beeinträchtigt“ beurteilt.

Die deutsche Version der SCALE basiert auf der englischen Originalversion (Fowler et al. Dev. Med. Child. Neurol. 2009;51:607-14) und wurde im Rahmen der Doktorarbeit von Frau Julia Balzer erstellt

© Copyright 2009 englische Originalversion 2.3 - University of California Los Angeles/ Orthopedic Hospital Center for Cerebral Palsy
© Copyright 2013 für die deutsche Version – Universitäts-Kinderspital Zürich – Eleonorenstiftung

Appendix 5: Study 2 - English SCALE version (original)

Please see next pages.

SCALE: Selective Control Assessment of the Lower Extremity Score Sheet

Date: _____ Patient's Name: _____ DOB: _____ GMFCS level: _____

Diagnosis: ☐ spastic diplegia ☐ spastic quadriplegia ☐ spastic hemiplegia R L other: _____

Grade	Left					Right				
	Hip	Knee	Ankle	STJ	Toes	Hip	Knee	Ankle	STJ	Toes
Normal (2 points)										
Impaired (1 point)										
Unable (0 points)										
Total Limb Score	L=		R=							

Resisted Synergy

knee extension with resisted limb extension										
dorsiflexion with resisted limb flexion										

Descriptors

hip flexion contracture										
adductor contracture or spasticity										
knee flexion contracture										
hamstring tightness										
plantar flexion contracture										
plantar flexor spasticity										
inverts or everts, not pure dorsiflexion										
primarily moves toes										
mirrors motion on opposite limb										
motion slower than 3 second verbal count										
moves one direction only (note motion achieved)										
movement of other joints										
motion \leq 50% of available ROM										

Other comments regarding test: _____

Examiner

SCALE: Directions for administration

The patient must be able to follow simple motor commands. To test this ability, ask the patient to move his or her least affected body part. Before asking the patient to perform each joint test, passively move the joint to assess ROM. To assure understanding, demonstrate the movement sequence while supporting the limb. The language in the instructions to the patient is suggested and may be modified as needed to elicit optimum performance for individual patients. To guide patients in the desired speed of movement, provide a verbal three-second count during the task. Multiple attempts are allowed and feedback to improve performance is acceptable.

General instructions to patient – *"I am going to ask you to move in a certain way. Move the way I ask you to move. Try not to move any other part of your body. If you have any questions or you don't understand what I am asking you to do, please tell me."*

Hip

Position – Side lying with the hip and knee fully extended. Support the limb medially at the knee and ankle. For stability, you may flex the lower untested limb. The tested motion is hip flexion while keeping the knee extended. Assess hip flexion ROM with the knee extended, as it may be limited by hamstring tightness. If the patient has difficulty with this task because of hamstring tightness, then ask him or her to extend, flex then extend the hip while keeping the knee flexed 90°. Evaluate hip extension ROM to assure an adequate arc of motion to assess performance of the task.

Instructions to patient – *Ask the patient to flex, extend then flex the hip while keeping the knee extended. For example: "Move your leg forward, back then forward again while keeping your knee straight. I will take you through the motion first, and then I'd like you to do it yourself."*

Knee

Position – The remaining tests are done in sitting with legs over the edge of the exam table. During the remaining tests you may allow the patient to lean back on his or her hands so the trunk is approximately 20° from vertical to compensate for hamstring tightness.

Instructions – *Ask the patient to extend, flex then extend the knee while keeping the hip flexed. For example: "Straighten your knee as much as you can, then bend it and straighten again. Try to do this without leaning further back or moving your other leg. I will take you through the motion first, and then I'd like you to do it yourself."*

Limb Extension Synergy – If quadriceps weakness is suspected, limb extension synergy may be assessed. Allow the patient to lean back on his or her hands or be supported so the trunk is approximately 45° from vertical. Position the limb in hip and knee flexion with ankle dorsiflexion. Ask the patient to push against your hand, extending the knee and plantar flexing the foot and toes. Resist at the metatarsal heads and compare knee extension excursion to the amount achieved during the knee selective voluntary motor control test.

Ankle

Position – Sitting, as in the knee extension test. The knee is extended and the examiner supports the calf. Assess passive ankle dorsiflexion ROM with the knee extended. The knee may be flexed to approximately 20° if needed to accommodate hamstring and/or gastrocnemius tightness.

Instructions to patient – *Ask patient to dorsiflex, plantar flex then dorsiflex the ankle while maintaining knee extension. For example: "Keeping your knee straight while I support your leg, move your foot up, down then up again. I will take you through the motion first, then I'd like you to do it yourself."*

Limb Flexion Synergy (Confusion Test) – If dorsiflexor muscle weakness is suspected, limb flexion synergy may be assessed. Ask the patient to flex the hip while keeping the knee flexed. Resist hip flexion at the distal thigh. Compare dorsiflexion excursion to the amount achieved during the ankle selective voluntary motor control test.

Foot/Subtalar Joint

Position – Sitting, as in the knee and ankle tests. The calf is supported.

Instructions to patient – *Ask patient to invert, evert then invert while maintaining knee extension. For example: "Move your ankle in, then out then in again while I support your leg. I will take you through the motion first, then I'd like you to do it yourself."*

Toes

Position – Sitting, as in the ankle test. The heel is supported.

Instructions to patient – *Ask patient to flex, extend then flex toes without moving ankle or knee. For example: "Curl all your toes down, then up then down again while I support your leg. I will take you through the motion first, then I'd like you to do it yourself."*

SCALE: Selective Control Assessment of the Lower Extremity

Instructions for Grading

Each joint is scored either 2, 1 or 0 points. These are summed for a Total Limb Score. The number of points for each grade is in parentheses. For each joint, check the joint score and all applicable descriptors on the *SCALE Score Sheet*.

Hip

- Normal (2) Flexes, extends then flexes again. During flexion, movement occurs without knee flexion, within a three-second verbal count and without mirror movement (the same movement on the contralateral limb). If alternate hip extension test is used, extends, flexes then extends again. During extension, movement occurs without knee extension, within a three-second verbal count and without mirror movement.
- Impaired (1) One or more of the following occur: extends or flexes $\leq 50\%$ of available range of motion in the test position, performs task slower than three-second verbal count, exhibits mirror movements, movement occurs in only one direction or motion at untested joint occurs.
- Unable (0) Does not flex or extend hip or does so only with simultaneous knee movement.

Knee

- Normal (2) Extends, flexes and extends again. Movement occurs within three-second verbal count, without motion of the trunk or other joints and without mirror movement. A Normal grade may be given if the knee extends $> 50\%$ of available range of motion in the test position.
- Impaired (1) One or more of the following occur: extends $\leq 50\%$ of available range of motion, performs task slower than three-second verbal count, exhibits mirror movements, movement occurs in only one direction or motion at untested joint occurs.
- Unable (0) Does not extend or only extends with simultaneous hip or ankle movement.

Ankle

- Normal (2) Dorsiflexes, plantar flexes and dorsiflexes again. Movement occurs within a three-second verbal count, without motion at other joints and without mirror movement. At least 15° of ankle motion in the sagittal plane must be observed.
- Impaired (1) One or more of the following occur: dorsiflexes $\leq 50\%$ of available passive range of motion in the test position or active range during Limb Flexion Synergy, performs task slower than three-second verbal count, exhibits mirror movements, movement occurs in only one direction or motion at untested joint occurs. An "Impaired" grade is given if the motion is accompanied by toe extension or ankle inversion.
- Unable (0) Does not dorsiflex or only dorsiflexes with hip and knee flexion.

Foot/Subtalar Joint

- Normal (2) Inverts, everts and inverts again. Movement occurs within a three-second verbal count, without motion at other joints and without mirror movement. Active eversion must occur.
- Impaired (1) One or more of the following occur: inverts or everts $\leq 50\%$ of available range of motion, performs task slower than three-second verbal count, exhibits mirror movements, movement occurs in only one direction or motion at untested joint occurs.
- Unable (0) Does not invert or evert or movement occurs only in synergy pattern. May dorsiflex, plantar flex or not move ankle at all.

Toes

- Normal (2) Flexes, extends and flexes again. Movement occurs within a three-second verbal count, without motion at other joints and without mirror movement. Motion should occur at all five toes.
- Impaired (1) One or more of the following occur: flexes or extends $\leq 50\%$ of available range of motion, performs task slower than three-second verbal count, exhibits mirror movements, movement occurs in only one direction or motion at untested joint occurs.
- Unable (0) Does not flex or extend toes.

Difference between Unable and Impaired

Unable (total synergy) has simultaneous movement at two or more joints. For every degree of motion at the desired joint, concomitant obligatory motion that is a part of the synergy pattern occurs at another joint in the limb. Patients with impaired motor control may be able to move the desired joint through a small arc of motion without any other joint motion, however a portion of the motion is accompanied by motion at an adjacent joint.

Difference between Impaired and Normal

Normal motor control is the ability to isolate joint motion through more than 50% of the available ROM within a three-second verbal count in an alternating fashion. The motion occurs without accompanying motion at any other joints of either limb. The inability to perform this task is impaired.

Appendix 6: Study 3 - Search terms for all databases

Systematic Review evaluating psychometric studies of SVMC measures for the lower extremity in children with UMN lesions

Pubmed

((selective voluntary motor control[TIAB] OR selective movement*[TIAB] OR Selectivity[TIAB] OR voluntary movement*[TIAB] OR volitional movement*[TIAB] OR quality of movement*[TIAB] OR muscle activation[TIAB] OR muscular activation[TIAB] OR movement pattern*[TIAB] OR physiological movement*[TIAB] OR sensorimotor control[TW] OR foot targeting[TW] OR foot placing[TW] OR coordinated movement*[TIAB] OR motion compensation OR compensatory movement* OR muscular coactivation OR mirror movement* OR muscle synerg* OR muscular synerg* OR movement synerg* OR cortical overlap[TW] OR motor control[TW] OR joint coupling[TW] OR (interjoint OR inter-joint AND coordination) OR force modulation[TW] OR submaximal contraction*[TW] OR maximal voluntary contraction*[TW]) OR (cortical activation*[TW])) OR (selective motor control test*[TIAB]) OR (selective control assessment of the lower extremity [TIAB]) OR (foot-tapping-test[TIAB]) OR (lower extremity motor score[TIAB]) OR (gross motor performance measure[TIAB]) OR (quality function measure) OR (fugl meyer assessment) OR (brunnstrom recovery stages*[TIAB]) OR (stroke rehabilitation assessment of movement*[TIAB]) OR (observation) OR (lower extremity motor score) OR (brain motor control assessment) OR (voluntary response INDEX) OR (active range of motion*[TIAB]) OR (active ankle movement*[TIAB]) OR (electrogoniometer*[TIAB]) OR (video*[TIAB]) OR (electromyography*[TIAB]) OR (dynamometer*[TIAB]) OR (force transducer*[TIAB]) OR (torque transducer*[TIAB]) OR (relative phase) OR (target position*[TIAB]) OR (target force*[TIAB]) OR (target torque*[TIAB]) OR (target trajectory*[TIAB]) OR (visuomotor tracking) OR (tracking performance*[TIAB]) OR (maximal voluntary contraction) OR (submaximal torque steadiness*[TIAB]) OR (moment generation*[TIAB]) OR (reaction time) OR (isometric muscle force) OR (integrated electromyography) OR (muscle activation latenc*[TIAB]) OR (latenc* time*[TIAB]) OR (activation time*[TIAB]) OR (on-off ratio*[TIAB]) OR (activation ratio*[TIAB]) OR (reaction time frequency analysis*[TIAB]) OR (wavelet analysis*[TIAB]) OR (instantaneous mean frequency*[TIAB]) OR (principal component analysis*[TIAB]) OR (torque steadiness*[TIAB]) OR (accuracy INDEX*[TIAB]) OR (tracking INDEX*[TIAB]) OR (force INDEX*[TIAB]) OR (motor evoked potentials*[TIAB]) OR (transcranial stimulation*[TIAB]) OR (fmri*[TIAB]) OR (diffusion tensor imaging*[TIAB]) OR (fractional anisotropy*[TIAB])) OR ("Muscle, Skeletal"[MeSH] OR "Motor Skills Disorders"[Mesh] OR "Neuromuscular Manifestations"[Mesh] OR "Motor Activity"[Mesh] OR "Recovery of Function/physiology"[Mesh] OR "Gait Disorders, Neurologic"[Mesh])) AND (upper motor neuron lesion*[TW] OR "Cerebral Palsy"[Mesh] OR "Brain Injuries"[Mesh]) OR ("central nervous system diseases") OR ("Gait Disorders, Neurologic") OR ("pyramidal tract*") OR ("corticospinal tract*") AND ((instrumentation[sh] OR methods[sh] OR Validation Studies[pt] OR Comparative Study[pt] OR "psychometrics"[MeSH] OR psychometr*[tiab] OR clinimetr*[tw] OR clinometr*[tw] OR "outcome assessment (health care)"[MeSH] OR outcome assessment[tiab] OR outcome measure*[tw] OR "observer variation"[MeSH] OR observer variation[tiab] OR "Health Status Indicators"[Mesh] OR "reproducibility of results"[MeSH] OR reproducib*[tiab] OR "discriminant analysis"[MeSH] OR reliab*[tiab] OR unreliab*[tiab] OR valid*[tiab] OR coefficient[tiab] OR homogeneity[tiab] OR homogeneous[tiab] OR "internal

consistency"[tiab] OR (cronbach*[tiab] AND (alpha[tiab] OR alphas[tiab])) OR (item[tiab] AND (correlation*[tiab] OR selection*[tiab] OR reduction*[tiab])) OR agreement[tiab] OR precision[tiab] OR imprecision[tiab] OR "precise values"[tiab] OR test-retest[tiab] OR (test[tiab] AND retest[tiab]) OR (reliab*[tiab] AND (test[tiab] OR retest[tiab])) OR stability[tiab] OR interrater[tiab] OR inter-rater[tiab] OR intrarater[tiab] OR intra-rater[tiab] OR intertester[tiab] OR inter-tester[tiab] OR intratester[tiab] OR intra-tester[tiab] OR interobserver[tiab] OR inter-observer[tiab] OR intraobserver[tiab] OR intra-observer[tiab] OR intertechnician[tiab] OR inter-technician[tiab] OR intratechnician[tiab] OR intra-technician[tiab] OR interexaminer[tiab] OR inter-examiner[tiab] OR intraexaminer[tiab] OR intra-examiner[tiab] OR interassay[tiab] OR inter-assay[tiab] OR intraassay[tiab] OR intra-assay[tiab] OR interindividual[tiab] OR inter-individual[tiab] OR intraindividual[tiab] OR intra-individual[tiab] OR interparticipant[tiab] OR inter-participant[tiab] OR intraparticipant[tiab] OR intra-participant[tiab] OR kappa[tiab] OR kappa's[tiab] OR kappas[tiab] OR repeatab*[tiab] OR ((replicab*[tiab] OR repeated[tiab]) AND (measure[tiab] OR measures[tiab] OR findings[tiab] OR result[tiab] OR results[tiab] OR test[tiab] OR tests[tiab])) OR generaliza*[tiab] OR generalisa*[tiab] OR concordance[tiab] OR (intraclass[tiab] AND correlation*[tiab]) OR discriminative[tiab] OR "known group"[tiab] OR factor analysis[tiab] OR factor analyses[tiab] OR dimension*[tiab] OR subscale*[tiab] OR (multitrait[tiab] AND scaling[tiab] AND (analysis[tiab] OR analyses[tiab])) OR item discriminant[tiab] OR interscale correlation*[tiab] OR error[tiab] OR errors[tiab] OR "individual variability"[tiab] OR (variability[tiab] AND (analysis[tiab] OR values[tiab])) OR (uncertainty[tiab] AND (measurement[tiab] OR measuring[tiab])) OR "standard error of measurement"[tiab] OR sensitiv*[tiab] OR responsive*[tiab] OR ((minimal[tiab] OR minimally[tiab] OR clinical[tiab] OR clinically[tiab]) AND (important[tiab] OR significant[tiab] OR detectable[tiab]) AND (change[tiab] OR difference[tiab])) OR (small*[tiab] AND (real[tiab] OR detectable[tiab]) AND (change[tiab] OR difference[tiab])) OR meaningful change[tiab] OR "ceiling effect"[tiab] OR "floor effect"[tiab] OR "Item response model"[tiab] OR IRT[tiab] OR Rasch[tiab] OR "Differential item functioning"[tiab] OR DIF[tiab] OR "computer adaptive testing"[tiab] OR "item bank"[tiab] OR "cross-cultural equivalence"[tiab])))) AND Humans[Mesh])) NOT ("Randomized Controlled Trial"[Publication Type] OR "Controlled Clinical Trial"[Publication Type] OR "addresses"[Publication Type] OR "biography"[Publication Type] OR "case reports"[Publication Type] OR "comment"[Publication Type] OR "directory"[Publication Type] OR "editorial"[Publication Type] OR "festschrift"[Publication Type] OR "interview"[Publication Type] OR "lectures"[Publication Type] OR "legal cases"[Publication Type] OR "legislation"[Publication Type] OR "letter"[Publication Type] OR "news"[Publication Type] OR "newspaper article"[Publication Type] OR "patient education handout"[Publication Type] OR "popular works"[Publication Type] OR "congresses"[Publication Type] OR "consensus development conference"[Publication Type] OR "consensus development conference, nih"[Publication Type] OR "practice guideline"[Publication Type] OR "intervention studies"[MESH])) AND Humans [Mesh] AND child [Mesh]

Ebscohost: CINAHL and PsycINFO

selective voluntary motor control OR selective movement* OR selectivity OR selective motorOR selective motor control OR voluntary movement* OR volitional movement* OR violate movement* OR quality of movement* OR muscle activation OR muscular activation OR movement pattern* OR physiological movement* OR sensorimotor control OR foot targeting OR foot placing OR coordinated movement* OR motion compensation OR compensatory movement* OR muscular co-activation OR co-contraction OR mirror movement*OR muscle synerg* OR muscularsynerg* OR movement synerg* OR synerg* OR

muscle activation pattern OR cortical overlap OR motor control OR joint coupling OR (inter joint OR inter-joint AND coordination) OR force modulation OR submaximal contraction* OR maximal voluntary contraction* OR cortical activation OR selective motor control test OR selective control assessment of the lower extremity OR foot-tapping-test OR lower extremity motor score OR gross motor performance measure OR quality function measure OR fuglmeier assessment OR brunnstrom recovery stages OR stroke rehabilitation assessment of movement OR observation OR lower extremity motor score OR brain motor control assessment OR voluntary response index OR active range of motion OR active ankle movement OR electrogoniometer OR video OR electromyography OR dynamometer OR force transducer OR torque transducer OR relative phase OR target position OR target force OR target torque OR target trajectory OR visuomotor tracking OR tracking performance OR maximal voluntary contraction OR submaximal torque steadiness OR moment generation OR reaction time OR isometric muscle force OR integrated electromyography OR muscle activation latenc* OR latenc* time OR activation time OR on-off ratio OR activation ratio OR reaction time frequency analysis OR wavelet analysis OR instantaneous mean frequency OR principal component analysis OR torque steadiness OR accuracy index OR tracking index OR force index OR motor evoked potentials OR transcranial stimulation OR fMRI OR diffusion tensor imaging OR fractional anisotropy AND (upper motor neuron lesion*) OR ("Cerebral Palsy") OR ("Brain Injuries") OR ("central nervous system diseases") OR ("Gait Disorders, Neurologic") OR ("pyramidal tract*") OR ("corticospinal tract*") AND ((MH "Research Measurement+") OR (MH "Outcomes Research") OR instrumentation* OR methods OR validation stud* OR comparative stud* OR psychometr* OR clinimetr* OR clinometr* OR (MH "Outcomes (Health Care)+") OR (MH "Treatment Outcomes+") OR (MH "outcome assessment") OR outcome assessment OR outcome measure* OR observer variation OR (MH "Health Status Indicators") OR (MH "reproducibility of results") OR reproducib* OR (MH "discriminant analysis") OR reliab* OR unreliab* OR valid* OR coefficient OR homogeneity OR homogeneous OR "internal consistency" OR (cronbach* AND (alpha OR alphas)) OR (item AND (correlation* OR selection* OR reduction*)) OR agreement OR precision OR imprecision OR "precise values" OR test-retest OR (test AND retest) OR (reliab* AND (test OR retest)) OR stability OR interrater OR inter-rater OR intrarater OR intra-rater OR intertester OR inter-tester OR intratester OR intra-tester OR interobserver OR inter-observer OR intraobserver OR intra-observer OR intertechnician OR inter-technician OR intratechnician OR intra-technician OR interexaminer OR inter-examiner OR intraexaminer OR intra-examiner OR interassay OR inter-assay OR intraassay OR intra-assay OR interindividual OR inter-individual OR intraindividual OR intra-individual OR interparticipant OR inter-participant OR intraparticipant OR intra-participant OR kappa OR kappa's OR kappas OR repeatab* OR ((replicab* OR repeated) AND (measure OR measures OR findings OR result OR results OR test OR tests)) OR generaliza* OR generalisa* OR concordance OR (intraclass AND correlation*) OR discriminative OR "known group" OR factor analysis OR factor analyses OR dimension* OR subscale* OR (multitrait AND scaling AND (analysis OR analyses)) OR item discriminant OR interscale correlation* OR error OR errors OR "individual variability" OR (variability AND (analysis OR values)) OR (uncertainty AND (measurement OR measuring)) OR "standard error of measurement" OR sensitiv* OR responsive* OR ((minimal OR minimally OR clinical OR clinically) AND (important OR significant OR detectable) AND (change OR difference)) OR (small* AND (real OR detectable) AND (change OR difference)) OR meaningful change OR "ceiling effect" OR "floor effect" OR "Item response model" OR IRT OR Rasch OR "Differential item functioning" OR DIF OR "computer adaptive testing" OR "item bank" OR "cross-cultural equivalence")

EMBASE

(selective voluntary motor control) OR (selective movement*) OR (selectivity) OR (voluntary movement*) OR (volitional movement*) OR (violate movement*) OR (quality of movement*) OR (muscle activation) OR (muscular activation) OR (movement pattern*) OR (physiological movement*) OR (sensorimotor control) OR (foot targeting) OR (foot placing) OR (coordinated movement*) OR (motion compensation) OR (compensatory movement*) OR (muscular co-activation) OR (co-contraction) OR (mirror movement*) OR (muscle synerg*) OR (muscular synerg*) OR (movement synerg*) OR (synerg*) OR (muscle activation pattern) OR (cortical overlap) OR (motor control) OR (joint coupling) OR (inter joint OR inter-joint AND coordination) OR (force modulation) OR (submaximal contraction*) OR (maximal voluntary contraction*) OR (cortical activation) OR (selective motor) OR (selective motor control test) OR (selective control assessment of the lower extremity) OR (foot-tapping-test) OR (lower extremity motor score) OR (gross motor performance measure) OR (quality function measure) OR (fuglmeyer assessment) OR (brunnstrom recovery stages) OR (stroke rehabilitation assessment of movement) OR (observation) OR (lower extremity motor score) OR (brain motor control assessment) OR (voluntary response index) OR (active range of motion) OR (active ankle movement) OR (electrogoniometer) OR (video) OR (electromyography) OR (dynamometer) OR (force transducer) OR (torque transducer) OR (relative phase) OR (target position) OR (target force) OR (target torque) OR (target trajectory) OR (visuomotor tracking) OR (tracking performance) OR (maximal voluntary contraction) OR (submaximal torque steadiness) OR (moment generation) OR (reaction time) OR (isometric muscle force) OR (integrated electromyography) OR (muscle activation latenc*) OR (latenc* time) OR (activation time) OR (on-off ratio) OR (activation ratio) OR (reaction time frequency analysis) OR (wavelet analysis) OR (instantaneous mean frequency) OR (principal component analysis) OR (torque steadiness OR accuracy index) OR (tracking index) OR (force index OR motor evoked potentials) OR (transcranial stimulation) OR (fMRI) OR (diffusion tensor imaging OR fractional anisotropy) AND (upper motor neuron lesion*) OR (cerebral palsy) OR (brain injuries) OR (stroke) OR (spinal cord injuries) OR (central nervous system diseases) OR (gait disorders) OR (pyramidal tract*) AND ('validation study'/exp OR psychometr* OR clinimetr* OR 'outcome assessment'/exp OR outcome assessment*:ab,ti OR 'comparative study'/de OR outcome measure* OR 'observer variation'/exp OR 'observer variation':ab,ti OR 'reproducibility'/exp OR reproducib*:ab,ti OR 'discriminant analysis'/de OR reliab*:ab,ti OR unreliab*:ab,ti OR valid*:ab,ti OR coefficient:ab,ti OR homogeneity:ab,ti OR homogeneous:ab,ti OR 'internal consistency':ab,ti OR (cronbach*:ab,ti AND (alpha:ab,ti OR alphas:ab,ti)) OR (item:ab,ti AND (correlation*:ab,ti OR selection*:ab,ti OR reduction*:ab,ti)) OR agreement:ab,ti OR precision:ab,ti OR imprecision:ab,ti OR 'precise values':ab,ti OR test-retest:ab,ti OR (test:ab,ti AND retest:ab,ti) OR (reliab*:ab,ti AND (test:ab,ti OR retest:ab,ti)) OR stability:ab,ti OR interrater:ab,ti OR inter-rater:ab,ti OR intrarater:ab,ti OR intra-rater:ab,ti OR intertester:ab,ti OR inter-tester:ab,ti OR intratester:ab,ti OR intra-tester:ab,ti OR interobserver:ab,ti OR inter-observer:ab,ti OR intraobserver:ab,ti OR intra-observer:ab,ti OR intertechnician:ab,ti OR inter-technician:ab,ti OR intratechnician:ab,ti OR intra-technician:ab,ti OR interexaminer:ab,ti OR inter-examiner:ab,ti OR intraexaminer:ab,ti OR intra-examiner:ab,ti OR interassay:ab,ti OR inter-assay:ab,ti OR intraassay:ab,ti OR intra-assay:ab,ti OR interindividual:ab,ti OR inter-individual:ab,ti OR intraindividual:ab,ti OR intra-individual:ab,ti OR interparticipant:ab,ti OR inter-participant:ab,ti OR intraparticipant:ab,ti OR intra-participant:ab,ti OR kappa:ab,ti OR kappa*:ab,ti OR kappas:ab,ti OR repeatab*:ab,ti OR ((replicab*:ab,ti OR repeated:ab,ti) AND (measure:ab,ti OR measures:ab,ti OR findings:ab,ti OR result:ab,ti OR results:ab,ti OR test:ab,ti OR tests:ab,ti)) OR generaliza*:ab,ti OR generalisa*:ab,ti OR concordance:ab,ti OR (intraclass:ab,ti AND

correlation*:ab,ti) OR discriminative:ab,ti OR 'known group':ab,ti OR factor analysis:ab,ti OR factor analyses:ab,ti OR dimension*:ab,ti OR subscale*:ab,ti OR (multitrait:ab,ti AND scaling:ab,ti AND (analysis:ab,ti OR analyses:ab,ti)) OR item discriminant:ab,ti OR interscale correlation*:ab,ti OR error:ab,ti OR errors:ab,ti OR 'individual variability':ab,ti OR (variability:ab,ti AND (analysis:ab,ti OR values:ab,ti)) OR (uncertainty:ab,ti AND (measurement:ab,ti OR measuring:ab,ti)) OR 'standard error of measurement':ab,ti OR sensitiv*:ab,ti OR responsive*:ab,ti OR ((minimal:ab,ti OR minimally:ab,ti OR clinical:ab,ti OR clinically:ab,ti) AND (important:ab,ti OR significant:ab,ti OR detectable:ab,ti) AND (change:ab,ti OR difference:ab,ti)) OR (small*:ab,ti AND (real:ab,ti OR detectable:ab,ti) AND (change:ab,ti OR difference:ab,ti)) OR meaningful change:ab,ti OR 'ceiling effect':ab,ti OR 'floor effect':ab,ti OR 'Item response model':ab,ti OR IRT:ab,ti OR Rasch:ab,ti OR 'Differential item functioning':ab,ti OR DIF:ab,ti OR 'computer adaptive testing':ab,ti OR 'item bank':ab,ti OR 'cross-cultural equivalence':ab,ti)

SCOPUS

TITLE-ABS-KEY((selective voluntary motor control) OR (selective movement*) OR (selectivity) OR (selective motor) OR (selective motor control) OR (voluntary movement*) OR (volitional movement*) OR (violate movement*) OR (quality of movement) OR (muscle activation) OR (muscular activation) OR (movement pattern*) OR (physiological movement*) OR (sensorimotor control) OR (foot targeting) OR (foot placing) OR (coordinated movement*) OR (motion compensation) OR (compensatory movement*) OR (muscular co-activation) OR (co-contraction) OR (mirror movement*) OR (muscle synerg*) OR (muscular synerg*) OR (movement synerg*) OR (synerg*) OR (muscle activation pattern) OR (cortical overlap) OR (motor control) OR (joint coupling) OR (inter joint OR inter-joint AND coordination) OR (force modulation) OR (submaximal contraction*) OR (maximal voluntary contraction*) OR (cortical activation) OR (selective motor control test) OR (selective control assessment of the lower extremity) OR (foot-tapping-test) OR (lower extremity motor score) OR (gross motor performance measure) OR (quality function measure) OR (fugl meyer assessment) OR (brunnstrom recovery stages) OR (stroke rehabilitation assessment of movement) OR (observation) OR (lower extremity motor score) OR (brain motor control assessment) OR (voluntary response INDEX) OR (active range of motion) OR (active ankle movement) OR (electrogoniometer) OR (video) OR (electromyography) OR (dynamometer) OR (force transducer) OR (torque transducer) OR (relative phase) OR (target position) OR (target force) OR (target torque) OR (target trajectory) OR (visuomotor tracking) OR (tracking performance) OR (maximal voluntary contraction) OR (submaximal torque steadiness) OR (moment generation) OR (reaction time) OR (isometric muscle force) OR (integrated electromyography) OR (muscle activation latenc*) OR (latenc* time) OR (activation time) OR (on-off ratio) OR (activation ratio) OR (reaction time frequency analysis) OR (wavelet analysis) OR (instantaneous mean frequency) OR (principal component analysis) OR (torque steadiness) OR (accuracy INDEX) OR (tracking INDEX) OR (force INDEX) OR (motor evoked potentials) OR (transcranial stimulation) OR (fmri) OR (diffusion tensor imaging) OR (fractional anisotropy)) AND TITLE-ABS-KEY((upper motor neuron lesion*) OR (cerebral palsy) OR (brain injuries) OR (pyramidal tract*) OR (corticospinal tract*)) AND NOT (stroke) OR (adult*) AND ((MH "Research Measurement+") OR (MH "Outcomes Research") OR instrumentation* OR methods OR validation stud* OR comparative stud* OR psychometr* OR clinimetr* OR clinometr* OR (MH "Outcomes (Health Care)+") OR (MH "Treatment Outcomes+") OR (MH "outcome assessment") OR outcome assessment OR outcome measure* OR observer variation OR (MH "Health Status

Indicators") OR (MH "reproducibility of results") OR reproducib* OR (MH "discriminant analysis") OR reliab* OR unreliab* OR valid* OR coefficient OR homogeneity OR homogeneous OR "internal consistency" OR (cronbach* AND (alpha OR alphas)) OR (item AND (correlation* OR selection* OR reduction*)) OR agreement OR precision OR imprecision OR "precise values" OR test-retest OR (test AND retest) OR (reliab* AND (test OR retest)) OR stability OR interrater OR inter-rater OR intrarater OR intra-rater OR intertester OR inter-tester OR intratester OR intra-tester OR interobserver OR inter-observer OR intraobserver OR intra-observer OR intertechnician OR inter-technician OR intratechnician OR intra-technician OR interexaminer OR inter-examiner OR intraexaminer OR intra-examiner OR interassay OR inter-assay OR intraassay OR intra-assay OR interindividual OR inter-individual OR intraindividual OR intra-individual OR interparticipant OR inter-participant OR intraparticipant OR intra-participant OR kappa OR kappa's OR kappas OR repeatab* OR ((replicab* OR repeated) AND TITLE-ABS-KEY((measure OR measures OR findings OR result OR results OR test OR tests)) OR generaliza* OR generalisa* OR concordance OR (intraclass AND correlation*) OR discriminative OR "known group" OR factor analysis OR factor analyses OR dimension* OR subscale* OR (multitrait AND scaling AND (analysis OR analyses)) OR item discriminant OR interscale correlation* OR error OR errors OR "individual variability" OR (variability AND (analysis OR values)) OR (uncertainty AND (measurement OR measuring)) OR "standard error of measurement" OR sensitiv* OR responsive* OR ((minimal OR minimally OR clinical OR clinically) AND (important OR significant OR detectable) AND (change OR difference)) OR (small* AND (real OR detectable) AND (change OR difference)) OR meaningful change OR "ceiling effect" OR "floor effect" OR "Item response model" OR IRT OR Rasch OR "Differential item functioning" OR DIF OR "computer adaptive testing" OR "item bank" OR "cross-cultural equivalence")) AND (LIMIT-TO(PUBYEAR,2017) OR LIMIT-TO(PUBYEAR,2016) OR LIMIT-TO(PUBYEAR,2015)) AND (EXCLUDE(SUBJAREA,"AGRI") OR EXCLUDE(SUBJAREA,"BIOC") OR EXCLUDE(SUBJAREA,"BUSI") OR EXCLUDE(SUBJAREA,"CENG")) AND (EXCLUDE(SUBJAREA,"CHEM") OR EXCLUDE(SUBJAREA,"DECI") OR EXCLUDE(SUBJAREA,"DENT")) AND (EXCLUDE(SUBJAREA,"COMP") OR EXCLUDE(SUBJAREA,"EART") OR EXCLUDE(SUBJAREA,"ECON") OR EXCLUDE(SUBJAREA,"ENER")) AND (EXCLUDE(SUBJAREA,"PSYC") OR EXCLUDE(SUBJAREA,"PHAR") OR EXCLUDE(SUBJAREA,"IMMU")) AND (LIMIT-TO(LANGUAGE,"English")) AND (EXCLUDE(EXACTKEYWORD,"Adult") OR EXCLUDE(EXACTKEYWORD,"Nonhuman"))

Cochrane Search Term

("selectivity" OR "selective motor control" OR "motor control" OR "torque" OR "muscle activity" OR "muscle activation") AND ("lower extremity" OR "leg") AND ("cerebral palsy" OR "acquired brain" OR "upper motor neuron lesion")

PEDRO

selective motor control

motor control AND cerebral palsy

Appendix 7: Study 3 - established specified COSMIN rating rules

Item	specified COSMIN rule
For all Boxes	
Item 01 'Was the percentage missing items described?' Item 02 'Was described how missing items were handled?'	As SVMC instruments are either assessments or laboratory-based tools, and not questionnaires, item 01 and 02 are not relevant, therefore these items should be scored as 'NA'. Nevertheless, completeness of the data reported in the paper should be checked and if any unexplained missing data are detected, this should be scored under the item 'methodological flaws'.
Item 03 'Sample size adequate?'	Omitting item 03 for the overall COSMIN box scores
Item 'Were there any important flaws in the design or methods of the study?'	<ul style="list-style-type: none"> - Unexplained missing data is detected - Sample characteristics of the comparator group (normally developed children) is missing/lacking
BoxB Reliability	
BoxB07 'Were patients stable...?' BoxB08 'Was the time interval appropriate?' BoxB09 'Were the test conditions similar?'	If more than one type of reliability is scored in one paper (i.e. inter-rater and test-retest reliability) Box B is filled out for each type of reliability separately (<i>as the items 07-09 will be scored differently</i>).
BoxB04 'Were at least two measurements available?' BoxB05 'Were the administrations independent?'	For inter-rater reliability testing of SVMC clinical assessments (i.e SCALE; Trost, SMC) parallel observation is possible, or rating via video scoring (<i>as these assessments evaluate SVMC via observation of the movement quality</i>). The scoring of the two administrations should be independent.
BoxB07 'Were patients stable...?'	SVMC stable patient = no interventions aimed at improving motor control, no changes in medications (which might have an influence on motor control) between the two test administrations
BoxB08 'Was the time interval appropriated?'	<p>Appropriate time interval for SVMC testing for:</p> <ul style="list-style-type: none"> - Interrater Reliability: for SVMC clinical assessments (i.e SCALE; Trost, SMC) parallel observation possible, or rating via video scoring) (<i>as theses assessments evaluate SVMC via observation of the movement quality</i>) - Test-retest Reliability: 1 day until 8 weeks (<i>as SVMC is thought to be a stable impairment, as long as the patient's condition is stable</i>) - Intrarater Reliability: ≤ 2 weeks (<i>to make sure, that the rater is not aware of their scores on the first administration</i>)

BoxF Hypotheses Testing	
BoxF04 'Hypotheses formulated are priori?' BoxF05 'Expected direction of correlation stated?' BoxF06 'Expected magnitude of correlation stated?'	<p>If the paper stated no explicit hypotheses (i.e. cohort studies) only the aim/purpose of the study), item boxF04 should be rated as <i>'Unclear what was expected'</i> or <i>'Hypotheses vague or not formulated but possible to deduce what was expected'</i> and not as <i>'NA'</i> as for construct validity a hypothesis is needed.</p> <p>Consecutive for boxF05 and 06</p> <p><i>'Expected direction/magnitude of the correlations or differences NOT stated'</i> should be scored.</p>
BoxF07 'Convergent validity adequate description of comparator' BoxF08 'Convergent validity: measurements properties adequately described in the population of interest?'	<p>In the cohort studies included in the review, validity of the used SVMC instrument was established by comparing SVMC of CP and normally developed children. Therefore, no comparator measurement was used, in this cases items boxF07 boxF08 should be scored as <i>'NA'</i>. If the sample characteristics of the 'comparator group' (normally developed children) are missing or incomplete this should be scored under item boxF09 'methodological flaws'</p>

Appendix 8: Study 3 - Quality Criteria for Measurement Properties

Property	Rating	Quality criteria
Content validity (BoxD)	+	A clear description is provided of the measurement aim, the target population, the concepts that are being measured, and the item selection AND target population and (investigators OR experts) were involved in item selection
	?	A clear description of above-mentioned aspects is lacking OR only target population involved OR doubtful design or method
	-	No target population involvement
Criterion validity (Box H)	+	Convincing arguments that gold standard is “gold” AND correlation with gold standard >0.70
	?	No convincing arguments that gold standard is “gold” OR doubtful design or method
	-	Correlation with gold standard <0.70, despite adequate design and method
Construct validity (BoxF)	+	Specific hypotheses were formulated AND at least 75% of the results are in accordance with these hypotheses
	?	Doubtful design or method (e.g., no hypotheses)
	-	Less than 75% of hypotheses were confirmed, despite adequate design and methods
Reliability (BoxB)	+	ICC or weighted Kappa>0.70;
	?	Doubtful design or method (e.g., time interval not mentioned)
	-	ICC or weighted Kappa <0.70, despite adequate design and method
Responsiveness (BoxI)	+	Correlation with an instrument measuring the same construct >0.50; OR at least 75% of the results are in accordance with the hypotheses OR AUC>0.70 AND correlations with related construct is higher than with unrelated constructs
	?	Solely correlation determined with unrelated constructs
	-	Correlation an instrument measuring the same construct <0.50; OR <75% of the results are in accordance with the hypotheses; OR AUC<0.70 AND correlations with related construct is higher than with unrelated constructs

Appendix 9: Study 3 - Evidence Level of Measurement Properties

Levels of Evidence for the Overall Quality of the Measurement Properties (Based on the Cochrane Back Review Group, Van Tulder et al. 2003)³⁰

Level	Rating	Criteria
Strong	+++ or ---	Consistent findings in multiple studies of good methodological quality OR in one study of excellent methodological quality
Moderate	++ or --	Consistent findings in multiple studies of fair methodological quality OR in one study of good methodological quality OR in one study of good methodological quality
Limited	+ or -	One study of fair methodological quality
Conflicting	±	Conflicting findings
Unknown	?	Only studies of poor methodological quality

+= positive results; ?=indeterminate results; -=negative results.


Appendix 10: Publication study 1

Please see next pages.

ORIGINAL ARTICLE



Influence of trunk control and lower extremity impairments on gait capacity in children with cerebral palsy

Julia Balzer^a, Petra Marsico^a, Elena Mitteregger^{a,b,c}, Marietta L. van der Linden^d, Thomas H. Mercer^d and Hubertus J. A. van Hedel^a 

^aPediatric Rehab Research Group, Rehabilitation Center Affoltern am Albis, University Children's Hospital Zurich, Affoltern am Albis, Switzerland;

^bRegional Group Zurich Foundation Cerebral Palsy (RGZ), Zurich, Switzerland; ^cInstitute for Physiotherapy, Zurich University of Applied Studies, ZHAW, Winterthur, Switzerland; ^dCentre for Health, Activity and Rehabilitation Research, Queen Margaret University, Edinburgh, Scotland

ABSTRACT

Purpose: We investigated the combined impact of trunk control and lower extremities impairments on predicting gait capacity in children with cerebral palsy (CP) and evaluated relationships between trunk control and lower extremities impairments.

Methods: Data of 52 children with CP [29 boys, mean age 11 years 9 months (± 4 years 6 months)] were included in this observational study. Gait capacity was measured by the "modified Time Up and Go test". Experienced therapists performed the "Modified Ashworth Scale", "Manual Muscle Test", the "Selective Control Assessment of the Lower Extremity", and the "Trunk Control Measurement Scale". We calculated Spearman correlations coefficients (ρ) and performed regression analyses.

Results: Trunk control was the strongest predictor ($b = -0.624$, $p < 0.001$) when explaining the variance of gait capacity and remained in the model together with spasticity ($R^2 = 0.67$). Muscle strength and selectivity correlated moderately to strongly with the trunk control and gait capacity (-0.68 to -0.78), but correlations for the spasticity were low ($\rho < -0.3$).

Conclusions: The interconnection between trunk control, leg muscle strength and selectivity for gait capacity in children with CP was shown. It indicates the significance of these impairments in gait assessment and, potentially, rehabilitation.

IMPLICATIONS FOR REHABILITATION

Trunk control was the strongest predictor for gait capacity in a regression model with lower extremity spasticity, muscle strength and selectivity and age as independent variables. Lower extremity muscle strength, selectivity, and trunk control explained a similar amount of gait capacity variance which is higher than that explained by lower extremity spasticity. Lower extremity muscle strength and selectivity correlated strongly with trunk control.

Therefore, we cautiously suggest that a combined trunk control and lower extremity training might be promising for improving gait capacity in children with CP (Gross Motor Function Classification System level I–III), which needed to be tested in future intervention-studies.

ARTICLE HISTORY

Received 1 March 2017
Revised 31 July 2017
Accepted 13 September 2017

KEYWORDS

Spasticity; muscle strength; selective voluntary motor control; regression analysis

Introduction

Gait rehabilitation is one of the primary goals within neuro-rehabilitation. The benefits of walking regarding daily mobility and general health-related issues (e.g., bone density and cardiovascular health) are well known for instance for children with cerebral palsy (CP) [1,2]. As the lower extremities and the trunk are involved during walking, knowledge about their individual and combined impact on gait is essential for developing optimal gait-training-strategies [3]. Studies exploring the impact of different lower extremity impairments such as spasticity, contractures, muscle weakness, or selective voluntary motor control (SVMC) on functional ambulation and gross motor function in children with CP found that a lack of muscle strength and SVMC were the strongest predictors [4–11]. Spasticity and a decreased range of motion (contractures) of the lower extremities were found to have smaller effects on gait [5,8,10]. However, generalization of these study results is limited due to their methodological heterogeneity.

Traditionally, disturbed motor control and lower extremity impairments were seen as the primary and secondary gait deviations, respectively [1]. Abnormality of trunk kinematics during walking was mostly considered as a compensatory gait deviation. Recently, this focus has been amended by an increasing number of studies investigating the trunk and upper limbs during walking [3,12,13]. This shift in focus coincided with the development of new trunk control measures for children with CP [13,14]. The first studies in this field provide increasing evidence that altered trunk control during gait in children with CP should not be considered solely as a compensation for gait deviation (due to altered lower extremity functioning), but should also be considered a direct aspect of gait deviation [3,12]. Additional papers concluded that interventions aiming to improve gross motor function in children with CP should also incorporate trunk control training [15–17].

However, until now, studies of ambulatory function have investigated the contribution of either lower extremity impairments

CONTACT Julia Balzer  Julia.Balzer@kispi.uzh.ch  Pediatric Rehab Research Group, Rehabilitation Center Affoltern am Albis, University Children's Hospital Zurich, Muhlebergstrasse 104, Affoltern am Albis CH-8910, Switzerland
© 2017 Informa UK Limited, trading as Taylor & Francis Group

[4–11] or that of trunk control [3,12,13], with no study assessing both lower extremity impairments and trunk control and evaluating their influence on gait. Accordingly, the primary aim of this study was to investigate the impact of both lower extremity and trunk control impairments on gait capacity (i.e., in this study, the time needed to perform the modified Timed Up and Go test or mTUG). Based on previous studies [4,8,9,11], we hypothesized that

(i) negative, moderate (correlation coefficient >0.5) relationships exist between gait capacity and lower extremity muscle strength, SVMC and trunk control and (ii) a positive, weak (correlation coefficient <0.5) relationship exists between gait capacity and lower extremity spasticity. Furthermore, we expected leg muscle strength and trunk control to be the strongest predictors for gait capacity.

In addition, we were interested in investigating how different lower extremity impairments are related to trunk control. Although knowledge about the interdependence between leg and trunk impairments is currently lacking we speculated, based on clinical reasoning, that positive, moderate correlations would be apparent between muscle strength and SVMC with trunk control and also that a negative, weak relationship between spasticity and trunk control would exist.

Methods

Participants

In- and out-patients of the "Rehabilitation Centre Affoltern am Albis, University Children's Hospital Zurich" were recruited by convenience sampling. Inclusion criteria were: diagnosis of spastic CP, age between 5 and 20 years, ability to walk (Gross Motor Function Classification System (GMFCS) level I–IV), and ability to follow simple instructions. Participants with additional movement disorders, with an unstable situation regarding their medication to regulate muscle tone and/or who had a botulinum toxin injection within the last 6 months or any surgical correction within the last year were excluded. The study was approved by the ethical committee of the Canton of Zurich (KEK-ZH-Nr. 2011-0404). Informed consent and assent were obtained from parents and participants (respectively).

Measurements

All tests were carried out by the same two experienced neuro-pediatric physiotherapists within a maximum timeframe of 1 h and in accordance to standardized procedures.

Lower extremity assessments

Spasticity and muscle weakness of hip, knee, and ankle flexion and extension movements were assessed with the "Modified Ashworth Scale (MAS)," and the "Manual Muscle Test" (MMT), respectively.

The MAS [18] scores spasticity on an ordinal scale ranging from "0" to "4" in accordance to the velocity dependent definition of spasticity from Katz et al. [19]. Although its criterion validity was established by using the pendulum test [20], its correlation with an increased alpha-motor-neuron activation [21] as well as with increased muscle activation and resistance [22] ranged from weak to moderate only. In children with CP, interrater-reliability of the MAS for the lower extremity joints ranged from weak to good [22,23].

Muscle strength was evaluated with the MMT in accordance to Kendall [24]. Scores ranged from 0 to 5, while higher scores indicate higher muscle strength. Its scoring system was originally

developed and tested on validity for determining muscle weakness in patients with poliomyelitis [25]. Its interrater-reliability has not yet been tested in children with CP, but was moderate to good for children with muscular dystrophy [26]. Although evaluation of the psychometric properties for the MAS and MMT in children with CP is limited, we decided to perform them, as they are considered the clinical standard, and have also been used in previous studies, hence allow comparison of our results [5,8–11].

The "Selective Control Assessment of the Lower Extremity" (SCALE) assesses SVMC at the hip, knee, ankle, subtalar and toe joints and was specifically developed for children with CP. To evaluate the level of SVMC, the child is asked to perform specific and timed isolated movement patterns at each joint. Each joint movement is scored on a three point ordinal scale ranging from 0 to 2 (unable, impaired, normal). Its validity has been established by demonstrating strong correlations (Spearman's $\rho > 0.8$) with the GMFCS [27,28] and the Fugl-Meyer Test (items III–IV) [28] in children with spastic CP. Furthermore, a high level of interrater-reliability was demonstrated in this patient group (intraclass correlation coefficient (ICC) above 0.8) [27,28].

Trunk control assessment

Trunk control was assessed using the "Trunk Control Measurement Scale" (TCMS). This is a 15-item assessment that examines sitting balance during functional activities [14]. The TCMS takes into account that the trunk should provide a stable base of support and is also an actively moving body segment. The first five items test static sitting balance followed by 10 items testing dynamic sitting balance. Dynamic sitting balance is further divided into two subscales, seven items testing "selective movement control" and three items testing "dynamic reaching". Its validity was supported for children with spastic CP by (i) moderate to strong correlations with the "Gross Motor Function Measure" (GMFM) [14,29], (ii) significant differences between healthy children and children with CP [14], and (iii) a strong correlation with center of pressure measures while sitting [29]. Its interrater-reliability was established as the ICC was 0.91 [14].

Gait capacity assessment

For assessing the participants' gait capacity, the pediatric version of the mTUG [30] was performed. It records the time a child needs to stand up from a chair with foot contact, to walk three meter to a target, turn around and return to the chair and sit down. We performed two mTUG trials and calculated the average time needed. Reliability and validity of the mTUG were supported by a study in a sample of 176 children without physical disabilities and 41 young people with physical disabilities due to CP or spina bifida [30]. In our study, we performed the mTUG twice and included the average time of the two trials in our analyses.

Statistical analysis

Statistical analysis was performed with SPSS 17.0 (IBM, Armonk, NY). Alpha was set at 0.05 (two-tailed). The Shapiro-Wilk-test showed that the data of most scores were not normally distributed. Hence Spearman's correlation coefficients (ρ) were calculated between the mTUG, age, MMT, SCALE, MAS, and TCMS total and sub-scores. We also calculated ρ between the TCMS total and sub-scores and the MMT, SCALE, and MAS scores.

In a second step, simple and multiple linear regression analysis (backward modeling) were carried out to determine the most important predictor(s) for explaining mTUG variance. A model using MMT, SCALE, MAS, and TCMS total scores as independent

variables was analyzed. For the regression analysis, the following assumptions were checked (i) homogeneity of variance via a non-significant Levin's test; (ii) lack of multicollinearity, by calculating the tolerance and variance inflation factor for each independent variable; (iii) lack of autocorrelation, by calculating the Durbin-Watson test, and (iii) a lack of outliers (case-wise diagnostic) based on the values of Cook's and Mahalanobis' distance [31,32]. Based on the regression sample size guidelines by Miles and Shevlin [33] we aimed for a sample size of 50 participants which would be sufficient for a regression model with four predictors with a moderate to large effect size.

Results

Sixty-eight children with spastic CP gave informed consent for participation. Due to a lack of compliance (lack of motivation, concentration problems) or due to organizational issues (unavailable walking aids) data sets of 14 participants were incomplete. As case-wise diagnostic for the regression analysis revealed that mTUG scores of two participants (GMFCS level IV) laid three standard deviations above the mean, these participants were classed as outliers and omitted from the analyses. Therefore, demographic and performance characteristics of 52 participants are presented in Table 1. The 23 females and 29 males were on average 11 years and 9 months (SD 4 years 6 months) old. Twenty-two children had a GMFCS level I, 12 had level II, 16 level III, and two level IV. Further clinical characteristics are presented in Table 1.

Correlation analysis

Correlation results for the lower extremity impairments, TCMS, and mTUG are summarized in Table 2. The MMT total scores

showed the strongest relationship with both the mTUG and TCMS total score, closely followed by the correlations between the total SCALE scores and the mTUG and TCMS total score. Lowest correlations were found between the MAS total scores and the mTUG or TCMS total score and its sub-scores. Only the correlation between age and gait capacity was weak and non-significant. Corresponding scatter plots are shown in Figure 1. Furthermore, MMT and SCALE correlated strongly (Table 2).

Simple and multiple linear regression analysis

When applying simple linear regression modeling to predict gait capacity, the TCMS total score alone explained most of the variance (54%) of the mTUG, followed by the SCALE (43%), the MMT (40%), and the MAS (31%). As age was not correlated with the mTUG, it was not included in the regression analysis (Table 3).

We applied a multiple backward regression model to investigate which lower extremity and/or trunk impairments explain the greatest amount of variance in gait capacity. In the first step, the total SCALE score was removed from the model, followed by the MMT score. The TCMS was the strongest predictor with a standardized regression coefficient "b" of -0.624 ($p < 0.001$), when explaining the variance in mTUG. Together with the MAS, the TCMS remained in the final model and both explained overall 67% of the mTUG variance. To improve the interpretation of these findings, this analysis showed that a decrease of trunk control in the amount of 12 TCMS points resulted in a 6.6 s increase of the mTUG.

Discussion

The current study showed that trunk control appears to be the strongest predictor for gait capacity in children with CP and that leg muscle strength and SVMC are strongly related to trunk control in this group.

Prediction of gait capacity

Until now, no study has investigated the impact of both lower extremity impairments and trunk control on gait capacity. While we expected that trunk control and leg muscle strength were the strongest predictors for gait capacity, the MMT, to our surprise, and the SCALE scores were excluded from the regression model. The unanticipated exclusion of the MMT, as well as the exclusion of the SCALE, is likely to be caused by multicollinearity between TCMS, MMT, and SCALE. As multicollinearity is a methodological limitation of our study, its cause and consequences will be explained in further detail in the section below. The results of our simple regression analysis are in agreement with those of previous studies, which reported the importance of SVMC [4,6–8] and

Table 1. Participants' clinical and functional characteristics.

Measures	Spastic CP n = 52		
	Median	(IQR)	Range
MMT			
Total score (0–60)	44.0	(20)	20–60
SCALE			
Total score (0–20)	12.0	(5.2)	0–19
MAS			
Total score (0–48)	2.5	(4)	0–20
TCMS			
TCMS – static (0–20)	19.0	(5.2)	4–20
TCMS – selective (0–28)	14.0	(8.2)	1–26
TCMS – dynamic (0–10)	8.0	(4.2)	1–10
TMS – total (0–58)	41.0	(14.2)	6–56
mTUG (s)	7.9	(5.5)	4.3–47.5
Age (yy, mm)	11.7	(7.6)	5.9–19.11

MMT: Manual Muscle Test; SCALE: Selective Control Assessment of the Lower Extremity; MAS: modified Ashworth Scale; mTUG: modified Time Up and Go test; TCMS: Trunk Control Measurement Scale, SD: standard deviation; IQR: interquartile range.

Table 2. Spearman's correlation coefficients (q) between lower extremity impairments, trunk control and gait capacity.

Spearman's rank (q)	MMT	SCALE	MAS	mTUG
MMT (total)	1.00	0.849 ($p < 0.001$)	0.255 ($p = 0.068$)	0.787 ($p < 0.001$)
SCALE (total)	0.849 ($p < 0.001$)	1.00	0.435 ($p = 0.002$)	0.685 ($p < 0.001$)
MAS (total)	0.255 ($p = 0.068$)	0.435 ($p = 0.002$)	1.00	0.356 ($p = 0.010$)
TCMS – static	0.711 ($p < 0.001$)	0.604 ($p < 0.001$)	0.189 ($p < 0.001$)	0.695 ($p < 0.001$)
TCMS – selective	0.665 ($p < 0.001$)	0.717 ($p < 0.001$)	0.362 ($p < 0.001$)	0.493 ($p < 0.001$)
TCMS – dynamic	0.770 ($p < 0.001$)	0.675 ($p < 0.001$)	0.218 ($p < 0.001$)	0.614 ($p < 0.001$)
TCMS (total)	0.764 ($p < 0.001$)	0.757 ($p < 0.001$)	0.296 ($p = 0.033$)	0.597 ($p < 0.001$)
Age				0.093 ($p = 0.126$)

MMT: Manual Muscle Test; SCALE: Selective Control Assessment of the Lower Extremity; MAS: modified Ashworth Scale; mTUG: modified Time Up and Go test; TCMS: Trunk Control Measurement Scale.

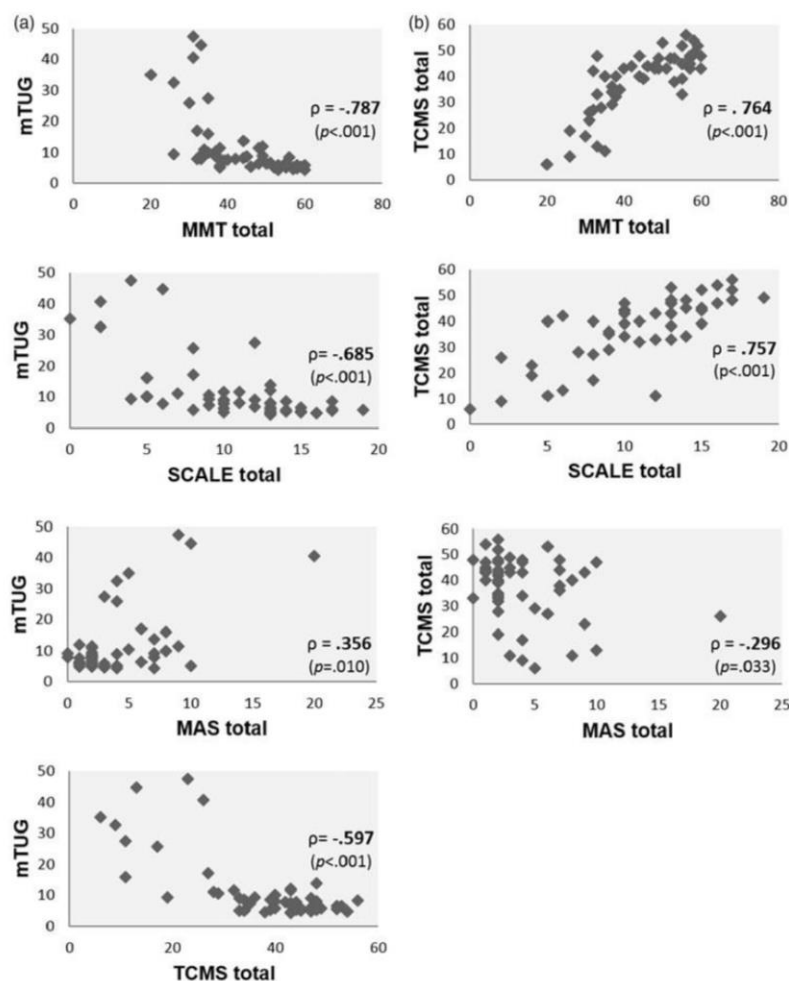


Figure 1. (a) Scatter plots and Spearman's correlation coefficients (ρ) between lower extremity impairments, trunk control and gait capacity; (b) scatter plots and Spearman's correlation coefficients (ρ) between lower extremity impairments and trunk control. MMT: Manual Muscle Test; SCALE: Selective Control Assessment of the Lower Extremity; MAS: modified Ashworth Scale; mTUG: modified Time Up and Go test; TCMS: Trunk Control Measurement Scale.

strength [6–8], and a minor influence of spasticity, on gait capacity/performance [6–8].

Nevertheless, a direct comparison regarding the absolute strength of the relationship, between our study and previously published research, is not appropriate due to the existence of several methodological differences: (i) previous studies used different dependent variables such as three dimensional gait analysis [10], or gross-motor function [5,9,11], (ii) differences in assessments/ methods were used to quantify lower extremity impairments, (iii) different levels of GMFCS of the study population, and (iv) different statistical analyses.

Comparing the simple and multiple regression results further in terms of the importance of trunk control on gait, we found only one other study which showed that trunk control (quantified by the "Segmental Assessment of Trunk Control") explained 38–40% variance of the GMFM in 92 children with CP (GMFCS I–V) [16]. In our study, the TCMS explained half of the variance of the mTUG within an ambulant sample (GMFCS I–IV). Although our results confirmed the strong relationship between trunk control

and gait capacity, a meaningful clinical interpretation of this finding in terms of causal relation between the two is difficult. This is due to the current lack of knowledge concerning the responsiveness of the TCMS and the lack of intervention studies which might have included the TCMS. Thereby it is unknown how likely it is to increase a patient's TCMS score and whether this results in an improvement of the mTUG.

Relationship between lower extremity motor functioning and trunk control

Regarding our secondary objective, this is, to our knowledge, the first study investigating the relationship between lower extremity impairments on trunk control in children with CP. Our a priori formulated hypotheses were confirmed by Spearman's rank correlation coefficients exceeding 0.7 for the MMT and SCALE with TCMS. However, the correlation between MAS and the TCMS was lower than expected. These outcomes seem to support two clinical impressions, formed prior to conducting this study, namely,

Table 3. Simple and multiple linear regression analysis for predicting gait capacity.

Dependent variable	Independent variable	B	Std. error B	b	R ²
mTUG	Simple linear regression				
	Constant	36.08	4.34		
	MMT (total score)	0.56	0.09	0.637 (p < 0.001)	0.40
	Constant	28.99	3.02		
	SCALE (total score)	-1.60	0.26	0.657 (p < 0.001)	0.43
	Constant	4.93	1.87		
	MAS (total score)	1.67	0.35	0.559 (p < 0.001)	0.31
	Constant	34.93	3.21		
	TCMS total	0.62	0.08	0.734 (p < 0.001)	0.54
	Constant	11.83	4.32		
	Age	0.04	0.36	0.016 (p ¼ 0.909)	0.00
	Multiple linear regression: MMT, SCALE, MAS, TCMS (backward modeling)				
	Step 1				
	Constant	26.91	4.34		
	SCALE	0.34	0.45	0.139 (p ¼ 0.459)	
	MMT	0.20	0.16	0.226 (p ¼ 0.213)	
	MAS	1.23	0.30	0.410 (p < 0.001)	
	TCMS	0.46	0.11	0.542 (p < 0.001)	0.68
	Step 2				
	Constant	28.88	3.81		
	MMT	0.12	0.12	0.136 (p ¼ 0.306)	
	MAS	1.11	0.26	0.371 (p < 0.001)	
	TCMS	0.44	0.11	0.520 (p < 0.001)	0.68 (p ¼ 0.459)
	Step 3				
	Constant	26.92	3.31		
	MAS	1.12	0.26	0.376 (p < 0.001)	
	TCMS	0.528	0.07	0.624 (p < 0.001)	0.67 (p ¼ 0.306)

MMT: Manual Muscle Test; SCALE: Selective Control Assessment of the Lower Extremity; MAS: modified Ashworth Scale; mTUG: modified Time Up and Go test; TCMS: Trunk Control Measurement Scale; B: Beta (unstandardized regression coefficient); Std. Error B: standardized error of Beta; b: standardized regression coefficient; R²: coefficient of determination.

that patients with better active trunk control (e.g., due to training) or passive trunk control (e.g., supported sitting or brace) have a better capacity for improving selective movements and strengthening of their lower extremities. Furthermore, the strong relationships between the trunk and the lower extremity functioning might be explained by their close neuroanatomical positions on Penfield's homunculus.

We identified only one recent study which addressed a similar topic. Heyrman et al. [12] investigated the impact of lower leg kinematics on trunk deviations in children with CP assessed during walking (as opposed to when sitting, as in our study). For measuring lower limb movements they used the Gait Profile Score (GPS). They found no significant correlations between the trunk parameters during gait (i.e., Trunk Profile Score) and the GPS ($r = 0.35$, $p = 0.13$) and only fair correlations between the TCMS and GPS ($r = -0.49$). Furthermore, the correlations between trunk parameters assessed in sitting (TCMS) and during gait were higher ($r = -0.63$ to -0.43). Therefore, they suggested that trunk deviations during walking are not exclusively associated with the presence of lower limb gait impairments and can thus be regarded as a discrete source of impairment and not merely a compensation [12].

Methodological considerations

As mentioned above, the problem of multicollinearity should be considered when interpreting these results [31]. The correlation matrix revealed high correlations ($q > 0.6$) between the MMT vs. SCALE, MMT vs. TCMS, and SCALE vs. TCMS. Furthermore, the average variance inflation factor of the starting model was above "1", which is considered as a threat to the validity of the model [30]. The presence of multicollinearity of the aforementioned variables makes it impossible to obtain unique estimates of the explained variance as these variables account for the similar

variance and their beta values are therefore interchangeable (Type II error) [34].

Our regression results showed that when predicting gait capacity by MMT, SCALE, MAS, and TCMS scores, the SCALE and MMT scores were removed from the model as their scores explained a similar amount of variance in mTUG variance as the TCMS. Note that these results do not indicate that SVMC and leg muscle strength do not influence gait capacity. The MAS, which on its own only correlates weakly with mTUG, seems to explain another part of the variance. Therefore, only the MAS and the variable with the highest beta value (TCMS) were kept in the final model.

This interpretation is supported by the results of the simple regression analyses, which showed that SCALE and MMT explained the second and the third largest amount of variance in mTUG (43% and 40%, respectively).

An alternative, if more complex approach to handle multicollinearity is to run a factor analysis (i.e., Structural Equation Modelling [6]) on the highly correlated predictors and to use the resulting factor scores (or latent variable) as a predictor [34]. As this statistical approach requires a larger sample size, it might be considered for future studies investigating similar research questions within a larger sample.

Concerning further methodological limitations about the generalizability of these study results, the dominance of participants with a higher gross motor/walking abilities level (GMFCS I and II ¼ 66.4% versus GMFCS III and IV ¼ 33.6%) should be considered. This underrepresentation of children with more severe mobility problems might also possibly explain the lower correlation with the MAS. The scatterplots of the MAS versus TCMS and mTUG reveal the dominance of participants with only a low level of spasticity.

Interpretation of our analysis is furthermore limited by the selected measures and joints/muscles. For instance, we only looked at flexors and extensors of the three major lower

extremity joints. We also did not record the participants' lower limb range of motion, which also can potentially affect a patient's gait capacity. However, previous studies have shown no or weak correlations of this impairment with gait or walking performance [10,11]. Furthermore, when interpreting our results related to muscle strength, it should be considered that we used an isometric strength test in order to allow for comparison with previous papers [5,8–11]. An isokinetic strength measure might have been more appropriated in relation to gait capacity.

Although it would have been of clinical interest to know if particular joints or other assessments explain more variance than other joints or assessments, this would have required a much larger sample size. Therefore, based on the sample size guidelines [33], we assigned only the total scores of the MAS, MMT, SCALE, and TCMS as predictors.

Another limitation of our study and those of previous studies [5,8–11] was that the chosen outcome measures (MAS, MMT, SCALE, and TCMS) are rated by the observer rather than "objectively" measured. To minimize a potential bias, the trained and experienced assessors (all of them were involved in previous studies on the translation, validation and reliability testing of these measures, see [28,29,35,36]) used standardized protocols and, except for the first author, were unaware of the hypotheses of this study.

Finally, it should be considered that the mTUG is a measure of both gait and balance activities, and thus might require more trunk and motor control than other commonly performed gait capacity tests (e.g., 0.10-meter walking test).

Clinical implications

First, the relatively weak correlations between the MAS and gait capacity, as indicated in the current study, showed that spasticity was not the main factor limiting (trunk control or) gait capacity in this sample of children with CP. This interpretation is supported by the previous studies [5,8–11] where spasticity, among other impairments (i.e., muscle weakness, impaired SVMC, contractures, learning difficulties) also did not have the strongest impact on gross motor function in children with CP. These findings as well as a recently increasing number of studies investigating the influence of SVMC on gait development [4] and gross motor function [5–7] challenge the traditionally claimed importance of spasticity management in ambulatory children with CP (GMFCS I–III).

Second, our results show, how, trunk control and lower extremity impairments independently and/or in combination may influence gait capacity. In addition we revealed, for the first time, the interrelationship of these two body functions. Based on our findings and those of previous studies investigating either trunk control [12–14,16] or gross motor performance [5,8–11], we cautiously suggest that therapists may wish to address the potential importance of trunk control as well as lower extremity functioning when attempting to improve gait capacity in children with GMFCS I–III.

Conclusions

The results of this study reveal that trunk control as well as lower extremity function, both assessed in sitting, are moderate to highly related to gait capacity in children with CP. Using a regression model to predict gait capacity, with lower extremity and trunk functioning as independent variables, we aimed to contribute new insight/knowledge for gait-rehabilitation in children with CP. Despite that the multiple regression models were limited by multicollinearity of some variables, we were able to show, on the

one hand, that trunk control, muscle strength, and SVMC account for a similar amount of gait capacity variance. Spasticity, on the other hand, accounts for the remaining but considerably lower amount of mTUG variance. This study provides also the first evidence that lower extremity strength, SVMC, and trunk control are highly correlated. Overall, the results of this study may indicate consideration of a case for combined strength and motor control training of the trunk and the lower extremity for gait-rehabilitation in children with CP (GMFCS I–III).

Acknowledgements

We thank all in- and out-patients and parents for their participation. We are grateful to the therapists of the "RGZ – Stiftung zugunsten cerebral Gelähmter", the "Vivendra-Stiftung" and the "Schule für Körper- und Mehrfachbehinderte (SKB) Zürich" for their assistance in recruiting the participants.

Disclosure statement

None of the funders were involved in the study design, data collection, analysis, and manuscript preparation and publication decisions. All ideas and decisions in relation to this study were made independently by the authors. No potential conflict of interest was reported by the authors.

Funding

This research project was funded by the "PhysioSwiss"; "Physiotherapy Science Foundation", "Mæxi-Foundation" and the "Swiss National Science Foundation (Project 32003B_156646)".

ORCID

Hubertus J. A. van Hedel  <http://orcid.org/0000-0002-9577-5049>

References

- [1] Gage JR, Schwartz MH, Koop SE, et al., editors. The identification and treatment of gait problems in cerebral palsy. London (UK): Mac Keith Press; 2009.
- [2] Bax M, Goldstein M, Rosenbaum P, et al. Proposed definition and classification of cerebral palsy, April 2005. *Dev Med Child Neurol*. 2005;47:571–576.
- [3] Saether R, Helbostad JL, Adde L, et al. The relationship between trunk control in sitting and during gait in children and adolescents with cerebral palsy. *Dev Med Child Neurol*. 2014;57:344–350.
- [4] Chruscikowski E, Fry NR, Noble JJ, et al. Selective motor control correlates with gait abnormality in children with cerebral palsy. *Gait Posture*. 2016;52:107–109.
- [5] Vos RC, Becher JG, Voorman JM, et al. Longitudinal association between gross motor capacity and neuromusculoskeletal function in children and youth with cerebral palsy. *Arch Phys Med Rehabil*. 2016;97:1329–1333.
- [6] Park E-Y, Kim W-H. Structural equation modeling of motor impairment, gross motor function, and the functional outcome in children with cerebral palsy. *Res Dev Disabil*. 2013;34:1731–1739.
- [7] Kim WH, Park EY. Causal relation between spasticity, strength, gross motor function, and functional outcome in

- children with cerebral palsy: a path analysis. *Dev Med Child Neurol.* 2011;53:68–73.
- [8] Ross S, Engsberg JR. Relationships between spasticity, strength, gait, and the GMFM-66 in persons with spastic diplegia cerebral palsy. *Arch Phys Med Rehabil.* 2007;88:1114–1120.
 - [9] Voorman JM, Dallmeijer AJ, Knol DL, et al. Prospective longitudinal study of gross motor function in children with cerebral palsy. *Arch Phys Med Rehabil.* 2007;88:871–876.
 - [10] Desloovere K, Molenaers G, Feys H, et al. Do dynamic and static clinical measurements correlate with gait analysis parameters in children with cerebral palsy? *Gait Posture.* 2006;24:302–313.
 - [11] Østensjø S, Carlberg EB, Vøllestad NK. Motor impairments in young children with cerebral palsy: relationship to gross motor function and everyday activities. *Dev Med Child Neurol.* 2004;9:580–589.
 - [12] Heyrman L, Feys H, Molenaers G, et al. Altered trunk movements during gait in children with spastic diplegia: compensatory or underlying trunk control deficit? *Res Dev Disabil.* 2014;35:2044–2052.
 - [13] Attias M, Bonnefoy-Mazure A, Lempereur M, et al. Trunk movements during gait in cerebral palsy. *Clin Biomech (Bristol, Avon).* 2014;30:28–32.
 - [14] Heyrman L, Molenaers G, Desloovere K, et al. A clinical tool to measure trunk control in children with cerebral palsy: the Trunk Control Measurement Scale. *Res Dev Disabil.* 2011;32:2624–2635.
 - [15] Desloovere K, Samson-fang L. Trunk control in children with cerebral palsy: where are we now? *Dev Med Child Neurol.* 2015;57:310–311.
 - [16] Curtis DJ, Butler P, Saavedra S, et al. The central role of trunk control in the gross motor function of children with cerebral palsy: a retrospective cross-sectional study. *Dev Med Child Neurol.* 2014;57:351–357.
 - [17] Rodby-Bousquet E, Czuba T, Hagglund G, et al. Postural asymmetries in young adults with cerebral palsy. *Dev Med Child Neurol.* 2013;55:1009–1015.
 - [18] Bohannon RW, Smith MB. Interrater reliability of a modified Ashworth scale of muscle spasticity. *Phys Ther.* 1987;67:206–207.
 - [19] Katz RT, Rovai GP, Brait C, et al. Objective quantification of spastic hypertonia: correlation with clinical findings. *Arch Phys Med Rehabil.* 1992;73:339–347.
 - [20] Naghdi S, Ebrahimi I, Asgari A, et al. A preliminary study into the criterion validity of the Modified Modified Ashworth Scale using the new measure of the alpha moto-neuron excitability in spastic hemiplegia. *Electromyogr Clin Neurophysiol.* 2007;47:187–192.
 - [21] Fleuren JF, Voerman GE, Erren-Wolters CV, et al. Stop using the Ashworth Scale for the assessment of spasticity. *J Neurol Neurosurg Psychiatry.* 1992;8:46–52.
 - [22] Fosang AL, Galea MP, McCoy AT, et al. Measures of muscle and joint performance in the lower limb of children with cerebral palsy. *Dev Med Child Neurol.* 2003;45:664–670.
 - [23] Mutlu A, Livanelioglu A, Gunel MK. Reliability of Ashworth and Modified Ashworth scales in children with spastic cerebral palsy. *BMC Musculoskelet Disord.* 2008;10:44.
 - [24] Kendall FP, McCreary EK, Provance PG. *Muscles: testing and function.* Baltimore (MD): Williams & Wilkins; 1993.
 - [25] Lovett RW, Martin EG. Certain aspects of infantile paralysis with a description of a method of muscle testing. *JAMA.* 1916;66:729–733.
 - [26] Cuthbert SC, Goodheart GJ. On the reliability and validity of manual muscle testing: a literature review. *Chiropr Osteopat.* 2007;15:4.
 - [27] Fowler EG, Staudt L, Greenberg MB, et al. Selective Control Assessment of the Lower Extremity (SCALE): development, validation, and interrater reliability of a clinical tool for patients with cerebral palsy. *Dev Med Child Neurol.* 2009;51:607–614.
 - [28] Balzer J, Marsico P, Mitteregger E, et al. Construct validity and reliability of the selective control assessment of the lower extremity in children with cerebral palsy. *Dev Med Child Neurol.* 2016;58:167–172.
 - [29] Mitteregger E, Marsico P, Balzer J, et al. Translation and construct validity of the Trunk Control Measurement Scale in children and youths with brain lesions. *Res Dev Disabil.* 2015;45–46:343–352.
 - [30] Williams EN, Carroll SG, Reddihough DS, et al. Investigation of the timed 'up & go' test in children. *Dev Med Child Neurol.* 2005;47:518–524.
 - [31] Bowerman B, O'Connell RT. *Linear statistical models: an applied approach.* 2nd ed. Belmont (CA): Duxbury Classic; 2000.
 - [32] Field A. *Discovering statistics using SPSS.* 2nd ed. London (UK): SAGE; 2005.
 - [33] Miles J, Shevlin M. *Applying regression and correlation. A guide for students and researchers.* London (UK): SAGE; 2001.
 - [34] Hucheson GD, Sofroniou N. *Multivariate social scientist introductory statistics using.* London (UK): SAGE; 1999.
 - [35] Marsico P, Mitteregger E, Balzer J, et al. Trunk Control Measurement Scale: reliability and discriminative validity in children and youths with neuro-motor disorders. *Dev Med Child Neurol.* 2017;59:699–705.
 - [36] Marsico P, Frontzek-Weps V, Balzer J, et al. Hypertonia Assessment Tool (HAT): reliability and validity in children with neuro-motor disorders. *J Child Neurol.* 2017;32:132–138.

Appendix 11: Publication study 2

Please see next page.

Construct validity and reliability of the Selective Control Assessment of the Lower Extremity in children with cerebral palsy

JULIA BALZER¹ | PETRA MARSICO¹ | ELENA MITTEREGGER² | MARIETTA L VAN DER LINDEN³ | THOMAS H MERCER³ | HUBERTUS J A VAN HEDEL¹

¹ Pediatric Rehab Research Group, Rehabilitation Center for Children and Adolescents, University Children's Hospital Zurich, Affoltern am Albis; ² Regional Group Zurich Foundation Cerebral Palsy (RGZ), Zurich, Switzerland. ³ Rehabilitation Sciences, School of Health Sciences, Queen Margaret University, Edinburgh, UK.

Correspondence to Julia Balzer, Pediatric Rehab Research Group, Rehabilitation Center for Children and Adolescents, University Children's Hospital Zurich, M€uhlebergstrasse 104, CH-8910 Affoltern am Albis, Switzerland. E-mail: Julia.Balzer@kispi.uzh.ch

PUBLICATION DATA

Accepted for publication 14th April 2015. Published online

ABBREVIATIONS

FMA	Fugl-Meyer Assessment
ICC	Intraclass correlation coefficient
IQR	Interquartile range
MAS	Modified Ashworth Scale
MDC	Minimal detectable change
MMT	Manual Muscle Test
SCALE	Selective Control Assessment of the Lower Extremity
SVMC	Selective voluntary motor control

AIM Assessing impaired selective voluntary movement control in children with cerebral palsy (CP) has gained increasing interest. We investigated construct validity and intra- and interrater reliability of the Selective Control Assessment of the Lower Extremity (SCALE). **METHOD** Thirty-nine children (21 males, 18 females) with spastic CP, mean age 12 years 6 months [range 6y 11mo–19y 9mo], Gross Motor Function Classification System (GMFCS) levels I to IV, participated. Differences in SCALE scores were determined on joint levels and between patients categorized according to their limb distribution and GMFCS levels. SCALE scores were correlated with the Fugl-Meyer Assessment, Manual Muscle Test, and Modified Ashworth Scale. To determine reliability, the SCALE was applied once and recorded on video.

RESULTS SCALE scores differed significantly between the less and more affected leg ($p < 0.001$) and between most leg joints. Total SCALE scores differed significantly between GMFCS levels I and II. Correlations with Fugl-Meyer Assessment, Manual Muscle Test, and Modified Ashworth Scale were 0.88, 0.88, and -0.55 respectively. Intraclass correlation coefficients were all above 0.9, with the minimal detectable change below 2 points.

INTERPRETATION The SCALE appears to be a valid and reliable tool to assess selective voluntary movement control of the legs in children with spastic CP.

With an incidence of 2 to 3 per 1000 in Europe, cerebral palsy (CP) is the most common motor disorder in childhood.¹ Depending on the severity and location of the congenital brain lesion, the appearance of positive and negative motor signs are heterogeneous.² Positive motor signs are associated with an involuntarily increased frequency or magnitude of muscle activity (i.e. hypertonia), whereas negative signs are characterized by insufficient muscle activity (i.e. muscle weakness) and their control (i.e. selective voluntary motor control [SVMC]).² Impaired SVMC has been defined as the inability 'to isolate the activation of muscles in a selected pattern in response to demands of a voluntary movement or posture'.² It is one of the most common motor impairments of the lower extremity in children with spastic CP.^{2,3} As impaired SVMC can be caused by a reduction of corticospinal drive as well as by increased input of descending subcortical pathways, consensus about its exact pathophysiological nature is still lacking.^{2–4} Although in comparison to other motor signs (e.g. hypertonia, muscle strength) improving SVMC has received little attention in the past decades,

recent studies have indicated the importance of SVMC in relation to motor performance.^{5–7} The results of several studies suggest that a loss of SVMC interferes much more with motor performance, such as walking, than, for instance, hypertonia and contractures.^{5–7} Furthermore, impaired selective activation can initiate and worsen a vicious cycle of limited active movement, joint contractures, hampered motor function, and diminished activity, thereby causing pain and appearance of secondary deformities in children with CP.⁸ Although the clinical importance of physiological muscle activation is obvious, routinely assessing selectivity is rare in the clinical environment, which, in turn, hampers evaluation of therapy-induced changes in SVMC.⁹

This lack of clinical assessment might be explained by the fact that testing SVMC is challenged by the coexistence of other motor signs. For instance, besides impaired SVMC, increased muscle tone or a lack of muscle strength, range of motion, sensory awareness, or stability in other joints can also result in limitations of movement quality.⁸ Therefore, requirements of measures evaluating SVMC are

high and only a few tools exist in the rehabilitation setting, especially for young patients with bilateral involvement.^{5,9} Recently, a clinical tool, the Selective Control Assessment of the Lower Extremity (SCALE), was developed to assess SVMC of the lower extremity in children with CP.¹⁰ In comparison to other SVMC tests,¹¹ the SCALE's testing procedure favourably attempts to differentiate between muscle weakness and a lack of selective control, by rating a joint movement as 'normal', when the patient is able to move his/her joint selectively (i.e. without co- and mirror movements and within the three verbal count [i.e. 'extend, flex, and extend again']), even if the patient can only perform 50% of the passive range of motion. Furthermore, the SCALE's check box allows for a more precise description of the nature of impaired selectively (i.e. contractures, mirror movements, co-movement on other joints), which can add important clinical information. Its content validity was approved by experienced clinicians and construct validity was excellent, as SCALE scores correlated well with the Gross Motor Function Classification System (GMFCS).¹⁰ The SCALE confirmed increasing proximal to distal impairment in children with CP, which can be explained by the somatotopic organization of the lower extremity in the sensorimotor cortex.¹² Furthermore, the SCALE's interrater reliability was excellent with intraclass correlation coefficients (ICCs) ranging from 0.88 to 0.91.¹⁰

Three recent studies,^{13–15} all from the SCALE developers, showed its predictive ability in relation to neurologically-induced gait disorders. Despite these interesting and promising findings, its construct validity has not been further investigated. Therefore, the aim of this study was to further evaluate validity of the SCALE in children with spastic CP. Based on results from previous studies, we hypothesized that SCALE scores would differ significantly between adjacent or contralateral joint pairs (i.e. lower in distal joints, lower in more affected limbs, respectively), less and more affected limbs, and GMFCS levels. To establish concurrent validity of the SCALE, a high positive correlation ($p > 0.70$) between the SCALE and the Fugl-Meyer Assessment (FMA) was expected. We further hypothesized that children with spastic CP and a high degree of muscle weakness and/or spasticity would score low via the SCALE assessment. Finally, we hypothesized that reliability would be excellent, with ICC values exceeding 0.8 (see also Fowler et al.¹⁰) and accompanied by acceptable levels of absolute measurement error.

METHOD

Participants

In- and outpatients of the Rehabilitation Centre Affoltern am Albis, University Children's Hospital Zurich were recruited by convenience sampling. A minimum sample size of 25 to 30 participants was required, in order to provide an accurate estimate of the random error.¹⁶ Inclusion criteria were: diagnosis of CP, aged between 5 and 20 years, ability to walk (GMFCS levels I–IV), and ability to

What this paper adds

- The Selective Control Assessment of the Lower Extremity (SCALE) German version is a valid and reliable assessment for children with spastic CP.
- Total SCALE scores differed significantly between Gross Motor Function Classification System levels I versus II.
- Minimal detectable change of the SCALE is 2 points.

follow simple instructions. Participants with an unstable situation regarding their tonus-regulating medications and/ or who had a botulinum toxin injection within the last 6 months, or any surgical correction within the last year, were excluded. The study was approved by the ethical committee of the Canton of Zurich (KEK-ZH-Nr.2011-0404). Informed consent and assent were obtained from parents and participants.

Measurements

In order to promote assessment of the SCALE in the German-speaking clinical environment, the SCALE was translated into German according to international guidelines:¹⁷

(1) translation into German by two independent native German-speaking physiotherapists; (2) creation of a consensus version; (3) back-translation into English by a translation company; and (4) endorsement by the authors of the original version (see Appendix S1, online supporting information, for the final German SCALE version).

Testing procedures were standardized according to the assessment guidelines. All tests were carried out by the same two experienced neuropaediatric physiotherapists, one assessing and one assisting. Tests were performed for both legs within a maximum time-frame of 1 hour.

SCALE administration required patients to perform specific isolated movement patterns at the hip, knee, ankle, subtalar, and toe joint. SVMC of each joint movement was scored on a 3-point ordinal scale. SVMC was scored as 'normal' (2 points) if the patient could move the tested joint isolated (e.g. without moving other joints), within at least 50% of the possible range of motion, and at a physiological cadence cued verbally by the therapist (e.g. 'flex, extend, flex'). If any deviation in performance occurred (movement performed slower, below 50% of range of movement, with co-/mirror-/synergistic-movements), selectivity was regarded as impaired (1 point). The score unable was given, if no joint movement could be made or mass-synergypatterns occurred. SVMC was scored separately for each joint, for each limb, and for both limbs together.

To analyze discriminant validity, patients were classified according to their limb involvement and GMFCS¹⁸ level (I–IV). The Manual Muscle Test (MMT) leg-score was used to determine the more and less affected leg. If MMT scores were similar, further differentiation was based on

Modified Ashworth Scale (MAS) scores.

To assess the SCALE's concurrent validity, the FMA¹⁹ was measured. The FMA is a valid assessment tool for testing SVMC in stroke and contains specific items for testing selectivity of the knee (FMA items IIIa; IVa) and ankle joint (FMA items IIIb; IVb). Like the SCALE, the FMA

uses a 3-point ordinal scale to score (0=cannot perform; 1=performs partly; 2=performs fully) selectivity of the joint movement.

Furthermore, when correctly applied, the MMT should also reflect the selective activation of a muscle (group). We therefore assessed strength of the hip and knee flexors and extensors, and of ankle dorsi- and plantar-flexors by the MMT (0–5; Kendall et al.²⁰).

Despite spasticity and SVMC being different constructs, spasticity can negatively influence SVMC and therefore we were interested in correlating the SCALE with MAS scores (0–4; Bohannon and Smith²¹). We assessed the MAS also for hip, knee, and ankle joints. The SCALE assessment was videotaped for (intra- and interrater) reliability testing, in order to minimize participants' strain. The camera was positioned in front of the participant. This position allowed observation of the tested joint movement and of possible compensatory and mirror movements of the contralateral limb, as well as of other body parts. Although an additional video from the sagittal plane may have allowed for a more accurate evaluation of the range of motion of the ankle and knee joint, none of the raters experienced difficulties in evaluating whether the movement exceeded 50% of the passive range of motion (one criterion that differentiates between normal or impaired SVMC) or not. For reliability testing the videotaped assessment was scored twice within a timeframe of 6 to 8 weeks after the first scoring. Rater(s) was (were) blinded to the results of the first scoring (intrarater) or results from the other rater (interrater).

Statistical analysis

The Shapiro–Wilk test showed that most scores were not normally distributed, hence non-parametric statistical tests were used.

Therefore, a Friedman test was performed to determine whether SCALE scores differed between joint pairs of each leg. Alpha was set at 0.05 (two-tailed). Post hoc differences between adjacent joints (i.e. hip vs knee), as well as between sum scores of the more and less involved leg, were determined with the Wilcoxon signed rank test (to adjust for multiple comparisons, alpha was set at 0.01). Differences in total SCALE scores for children categorized via GMFCS level were evaluated with the Kruskal–Wallis test. We defined a priori that we performed post hoc successive pairwise testing between adjacent GMFCS levels (e.g. level I vs II, level II vs III) with Mann–Whitney U tests (post hoc tests: alpha=0.025). To further evaluate the validity of the SCALE Spearman's rank correlation coefficients (ρ) between SCALE scores on joint, limb, and total levels and FMA, MMT, and MAS scores were calculated. Relative intra- and interrater reliability was evaluated by ICCs (two-way mixed model; type absolute agreement) and corresponding 95% confidence intervals were calculated for the less and more involved leg. Absolute reliability was determined by the standard error of measurement ($SEM=SD\sqrt{1-r}$) and the minimal detectable change ($MDC=SEM\sqrt{2.91.96}$).

Statistical analysis was performed with SPSS 17.0 (IBM, Armonk, NY, USA).

RESULTS

Forty-two children with spastic CP gave informed consent to participate in this study. One child did not complete the assessments because of a lack of compliance. As allocation of the more and less affected leg was not possible in two data sets, these data sets were omitted from all analyses. Therefore, demographic and clinical characteristics of 39 children with spastic CP (unilateral $n=20$; bilateral $n=19$) were available. The mean age was 12y 6mo (SD 3y 7mo). Eighteen children were female. Twenty-three children had a GMFCS level I, five had level II, eight level III and three level IV. Further characteristics are presented in Table I.

Discriminant validity

SCALE scores of contralateral joint pairs (i.e. knee vs knee) of the less affected leg were significantly higher compared to those of the more affected leg, with the exception of the hip joint, $p=0.157$ (Fig. 1a). SCALE scores were generally lower for distal compared to proximal joints for both legs, with the exception of the ankle versus toes for the less affected leg, and bilaterally for the subtalar joint on toes (Fig. 1a). SCALE limb scores were higher for the less affected limb (median=7; interquartile range [IQR]=0–10) compared to the more affected limb (median=5; IQR=0–9; $p<0.001$). When classifying participants in accordance with their diagnosis, statistically significant differences between the less and more affected limb were present for children with unilateral limb involvement (less affected: median=9; IQR=7–10, vs more affected: median=4.5; IQR=3–6; $p<0.001$) and bilateral involvement (less affected: median=5; IQR=2.5–6, vs more affected: median=6; IQR=3.5–7; $p=0.003$) (Fig. 1b).

Table I: Participants characteristics

Measures	Spastic CP (n=39)		
	Mean (SD)	Median (IQR)	Range
Scale			
Less affected leg (9/10)	6.6(2.8)	7.0 (4.0)	0–10
More affected leg (9/10)	4.5(2.0)	5.0 (3.0)	0–9
Total score (9/20)	11.2 (4.5)	13.0 (4.5)	0–19
FMA			
Less affected leg (9/8)	5.7(2.9)	8.0 (4.0)	0–8
More affected leg (9/8)	3.8(2.5)	5.0 (3.5)	0–8
Total score (9/16)	9.6(5.1)	11.0 (8.0)	0–16
MMT			
Less affected leg (9/30)	24.8 (6.1)	28.0 (11.0)	10–30
More affected leg (9/30)	20.9 (5.5)	23.0 (9.0)	9–30
Total score (9/60)	45.6 (11.0)	49.0 (17.5)	20–59
MAS			
Less affected leg (9/24)	1.9(3.1)	1.0 (2.5)	0–14
More affected leg (9/24)	3.4(3.2)	2.0 (3.0)	1–16
Total score (9/48)	5.4(6.0)	3.0 (5.0)	1–30

CP, cerebral palsy; SD, standard deviation; IQR, interquartile range; SCALE, Selective Control Assessment of the Lower Extremity; FMA, Fugl-Meyer Assessment; MMT, Manual Muscle Test; MAS, Modified Ashworth Scale.

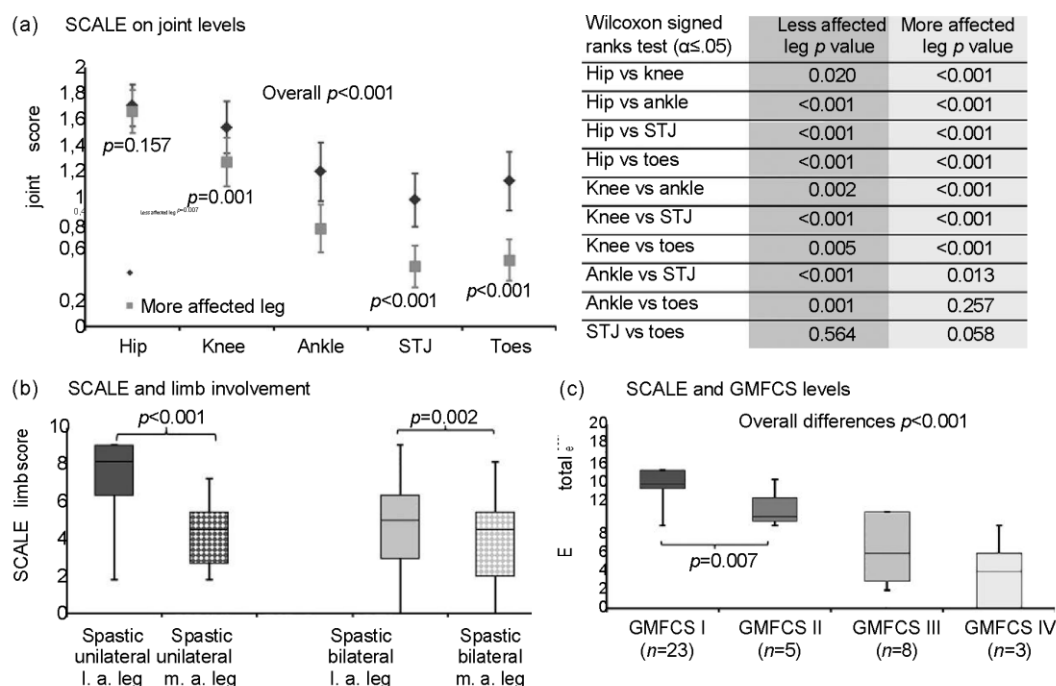


Figure 1: Discriminant validity for children with spastic cerebral palsy. (a) SCALE and joint pairs: SCALE joint scores between more and less involved limb and on adjacent pairs of joints: Friedman test and post hoc Wilcoxon signed rank test (post hoc tests: p values below 0.01 are considered significant). (b) SCALE and limb distribution: total SCALE scores of less affected (l.a.) versus more affected (m.a.) leg in children with spastic unilateral and bilateral limb involvement: Wilcoxon signed rank test. (c) SCALE and GMFCS levels: significant differences between all GMFCS levels (Kruskal–Wallis test) and between GMFCS I versus II (post hoc Mann–Whitney U test: p values below 0.025 are considered significant). SCALE, Selective Control Assessment of the Lower Extremity; STJ, subtalar joint; GMFCS, Gross Motor Function Classification System.

Furthermore, SCALE scores differed significantly between GMFCS levels ($p < 0.018$), and, more specifically, between GMFCS levels I and II (Fig. 1c).

Correlations

For the total SCALE score, high correlations²² between FMA and MMT were found (Fig. 2). The magnitude of the correlations between SCALE limb and joint scores, and the clinical measures were comparable to those presented for the total scores.

There was a negative moderate correlation between the SCALE and the MAS total scores.

Reliability

With ICC values exceeding 0.9 for limb and 0.8 for joint SCALE scores in children with spastic CP intra- and interrater reliability of the SCALE can be considered excellent. The MDC varied between 1.79 and 1.96 points (Table II).

DISCUSSION

Construct validity, as well as intra- and interrater reliability of the SCALE in children with spastic CP are supported by this study. Regarding the SCALE's discriminant validity

between adjacent or contralateral joint pairs, between more versus less affected limb, or between GMFCS levels, we could partly confirm our hypotheses in line with previous results.¹⁰

Concerning the SCALE's ability to discriminate between the more and less affected limb in children with hemi- and diplegia, differences were significant for both groups, but for the latter subgroup the difference was below the MDC.

On joint level, SCALE scores of the hip joint did not differ between the more and less impaired limb. This could be because of the limited number of participants with greater motor impairment at the hip (i.e. GMFCS level IV) and the large number of children with near maximal scores (ceiling effect, i.e. 10 participants had a maximum total SCALE score for their less affected leg). SCALE scores at most distal joint pairs tended to be lower, with the exception for comparison between the subtalar joint and the toes. This trend was observed previously by Fowler et al.¹² and Brunnstrom²³ who reported that selective inversion and eversion were described as the most challenging movements for children with CP, as well as adult stroke patients. As these movements rarely occur in isolation during daily activities (but frequently in combination

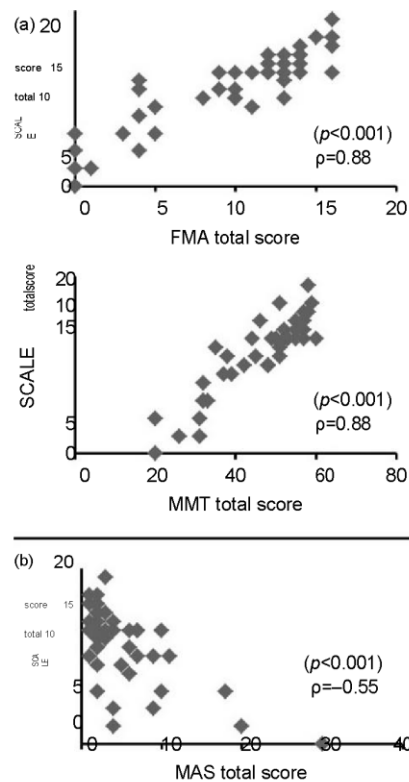


Figure 2: Spearman's rank correlation coefficient (ρ) of total Selective Control Assessment of the Lower Extremity (SCALE) scores and common clinical assessments for children with spastic cerebral palsy. (a) concurrent validity: SCALE versus Fugl-Meyer Assessment (FMA) and SCALE versus Manual Muscle Test (MMT). (b) correlation SCALE and Modified Ashworth Scale (MAS).

with the movement of other foot joints in supination or pronation) their movement performance might be experienced as unusual. Another neurophysiological explanation might be that the cortical representation of the lower extremity is largest for the big toe.²⁴

The SCALE's discriminant validity was reflected in an overall difference between the GMFCS levels. Nevertheless, because of the small sample size, we could only find significant differences between GMFCS levels I and II, and interpretation should be handled with caution. Performing a power analysis (80% power, two-tailed alpha 0.05) revealed that a sample size of 19 participants in GMFCS levels II and III, and 29 participants in GMFCS levels III and IV, would be required to determine statistically significant differences between these GMFCS levels.

The strong correlation between the SCALE and FMA, illustrated in Figure 2, confirms that both assessments measure broadly similar constructs. Correlation coefficients

Table II: Intra- and interrater reliability of the Selective Control Assessment of the Lower Extremity (SCALE)

	Spastic CP (n=38) ^a			
	Intrarater reliability		Interrater reliability	
	Less affected leg	More affected leg	Less affected leg	More affected leg
Descriptive				
Mean	6.55 (2.86)	4.63 (2.16)	6.55 (2.86)	4.63 (2.16)
(SD)1				
Mean	6.00 (2.81)	4.74 (2.50)	6.29 (2.94)	4.53 (2.16)
(SD)2				
Relative reliability				
ICC	0.95	0.96	0.94	0.91
p value	<0.001	<0.001	<0.001	<0.001
95% CI	0.90–0.97	0.93–0.98	0.89–0.97	0.84–0.96
Absolute reliability (SCALE points)				
SEM	0.71	0.64	0.69	0.68
MDC ₉₅	1.96	1.79	1.92	1.88

^aBecause of a failure in a video recording of the SCALE, we could include data from only 38 participants in the reliability analyses. CP, cerebral palsy; ICC, intraclass correlation coefficient; CI, confidence interval; SD, standard deviation; SEM, standard error of measurement; MDC₉₅, minimum detectable change at 95% confidence interval.

between SCALE and MMT were also of similar magnitude. This could indicate that a correctly applied MMT will partially reflect the ability to selectively activate a muscle (group).

Regarding our additional hypothesis in relation to the association between SVMC and spasticity, only a moderate negative correlation was found, with a large variation of SCALE scores in participants with low MAS scores (Fig. 2b). This range of SCALE scores in children with low spasticity might indicate that a mild level of spasticity does not necessarily affect SVMC negatively, while a more clear inverse relationship between SCALE and MAS is seen in participants with higher MAS values. However, the latter would have to be confirmed in studies including participants with a larger range of MAS values than reported in our study in which the majority were only mildly affected (i.e. mostly GMFCS levels I and II). In relation to the functional interdependence of SVMC with muscle weakness and spasticity, a possible influence of these impairments on the presented correlations cannot be excluded. Furthermore, like Fowler et al.,¹⁰ we found a high inverse relationship ($\rho \leq 0.80$) between the severity of CP (GMFCS levels) and the total SCALE score.

Our hypothesis, regarding reliability of the SCALE, was confirmed. We found excellent intra- and interrater reliability for the SCALE, as well as clinically acceptable values of absolute reliability for SCALE limb scores. As these results are based on a second rating of video recordings, the ICCs might be slightly higher than when rated via a second assessment, where interfering factors like the participant's compliance or state of health might have altered testing conditions. For future studies, accuracy of the video recordings could be improved by performing an additional video recording from the sagittal plane. However, our

current values are similar to previously reported observations.¹⁰ In relation to the absolute reliability of the SCALE, our study showed that an increase of more than 2 SCALE points for the more affected leg in children with CP could be considered a true change (MDC). A future study on the responsiveness of the SCALE will provide insight into whether such changes can be achieved with current rehabilitative (e.g. training or botulinum toxin) or surgical interventions (e.g. selective dorsal rhizotomy). In order to expand SCALE's application from a diagnostic tool to an evaluative one, and to counteract the limitation of its ordinal scale, it might be appropriate to consider combining its test procedure with a neurophysiological measure.^{4,25}

With regard to the methodological limitations of this study, we should mention that grouping the results for the less and more involved limb might have decreased variability between participants, which could have resulted in lower correlation coefficients and ICC values. Nevertheless, our observations are broadly comparable with previous reports.⁹

In conclusion, previous results about the SCALE's validity, interrater reliability, and increased distal impairment of SVMC were supported by this study. New evidence for construct validity of the SCALE in relation to common clinical tests in children with spastic CP, as well as important reliability aspects such as intrarater reliability and MDC values, were added.

ACKNOWLEDGEMENTS

We acknowledge contributions and approvals regarding the retranslated SCALE version from the authors of the original SCALE version (Eileen G Fowler, Loretta A Staudt, and Marcia B Greenberg, Department of Orthopedic Surgery, UCLA/Orthopedic Hospital Center for Cerebral Palsy, Los Angeles, USA). We thank all volunteer in- and outpatients and parents for their participation. We are grateful to the therapists of the following institutions: the RGZ – Stiftung zugunsten cerebraler Gelähmter; the Vivendra-Stiftung; and the 'Schule für Körper- und Mehrfachbehinderte' (SKB) Zürich for their assistance in recruiting the participants. This research project was funded by the PhysioSwiss, Switzerland; Physiotherapy Science Foundation, Switzerland; Maxi-Foundation, Switzerland; and the Swiss National Science Foundation (Project 32003B_156646), Switzerland. None of the funders were involved in the study design, data collection, analysis, and manuscript preparation and publication decisions. All ideas and decisions in relation to this study were made independently by the authors. The authors have stated that they had no interests which might be perceived as posing a conflict or bias.

SUPPORTING INFORMATION

The following additional material may be found online:

Appendix S1: Selective Control Assessment of the Lower Extremity (SCALE) – German version.

REFERENCES

1. SCPE. Surveillance of cerebral palsy in Europe: a collaboration of cerebral palsy surveys and registers. *Dev Med Child Neurol* 2000; 42: 816–24.
2. Sanger TD, Chen D, Delgado MR, Gaebler-Spira D, Hal-lest M, Mink JW. Definition and classification of negative motor signs in childhood. *Pediatrics* 2006; 118: 2159–67.
3. Bax M, Tydeman C, Flodmark O. Clinical and MRI correlated of cerebral palsy: the European Cerebral Palsy Study. *JAMA* 2006; 296: 1602–8.
4. Dobson F. Assessing selective motor control in children with cerebral palsy. *Dev Med Child Neurol* 2010; 52: 409–10.
5. Desloovere K, Molenaers G, Feys H, et al. Do dynamic and static clinical measurements correlate with gait analysis parameters in children with cerebral palsy? *Gait Posture* 2006; 24: 302–13.
6. Voorman JM, Dallmeijer AJ, Knol DL, Lankhorst GJ, Becher JG. Prospective longitudinal study of gross motor function in children with cerebral palsy. *Arch Phys Med Rehabil* 2007; 88: 871–6.
7. Østenjo S, Brogren E, Vollestad NK. Everyday functioning in young children with cerebral palsy: functional skills, caregiver assistance, and modifications of the environment. *Dev Med Child Neurol* 2003; 45: 603–17.
8. Miller F. *Physical Therapy of Cerebral Palsy*. New York: Springer, 2006.
9. Levin MF, Panturin E. Sensorimotor integration for functional recovery and the Bobath approach. *Mot Control* 2011; 15: 285–301.
10. Fowler EG, Staudt LA, Greenberg MB, Oppenheim WL. Selective Control Assessment of the Lower Extremity (SCALE): development, validation, and interrater reliability of a clinical tool for patients with cerebral palsy. *Dev Med Child Neurol* 2009; 51: 607–14.
11. Smits D-W, van Groenestijn AC, Ketelaar M, Scholtes VA, Becher JG, Gorter JW. Selective motor control of the lower extremities in children with cerebral palsy: inter-rater reliability of two tests. *Dev Neurorehabil* 2010; 13: 258–65.
12. Fowler E, Staudt L, Greenberg M. Lower-extremity selective voluntary motor control in patients with spastic cerebral palsy: increased distal motor impairment. *Dev Med Child Neurol* 2010; 52: 264–9.
13. Fowler E, Goldberg EJ. The effect of lower extremity selective voluntary motor control on interjoint coordination during gait in children with spastic diplegic cerebral palsy. *Gait Posture* 2009; 29: 102–7.
14. Goldberg EJ, Requejo PS, Fowler EG. Joint moment contributions to swing knee extension acceleration during gait in individuals with spastic diplegic cerebral palsy. *Gait Posture* 2011; 33: 66–70.
15. Goldberg EJ, Fowler EG, Oppenheim WL. Case reports: the influence of selective voluntary motor control on gait after hamstring lengthening surgery. *Clin Orthop Relat Res* 2012; 470: 1320–6.
16. Springate SD. The effect of sample size and bias on the reliability of estimates of error: a comparative study of Dahlberg's formula. *Eur J Orthod* 2012; 34: 158–63.
17. Beaton DE, Bobmardier C, Guillemin F, BosiFerraz M. Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine* 2011; 25: 3186–91.
18. Palisano RJ, Hanna SE, Rosenbaum PL, Tieman B. Probability of walking, wheeled mobility, and assisted mobility in children and adolescents with cerebral palsy. *Dev Med Child Neurol* 2010; 52: 66–71.
19. Danells CJ, Black SE. The Fugl-Meyer Assessment of motor recovery after stroke: a critical review of its measurement properties. *Neurorehabil Neural Repair* 2002; 16: 232–40.
20. Kendall F, Kendall McCreary E, Provance PG. *Muscle Testing and Function*. 4th edn. Philadelphia: Lippincott Williams & Wilkins, 1993.
21. Bohannon RW, Smith MB. Interrater reliability of a Modified Ashworth Scale of muscle spasticity. *Phys Ther* 1987; 67: 206–7.
22. Dawson-Saunders B, Trapp RG. *Basic & Clinical Biostatistics*. 2nd edn. Connecticut: Appleton & Lange, Prentice-Hall International, 1994.
23. Brunnstrom S. Motor testing procedures in hemiplegia: based on sequential recovery stages. *Phys Ther* 1966; 46: 357–75.
24. Rothwell JC. *Control of Human Voluntary Movement*. London: Croom Helm, 1987.
25. Zwaan E, Becher JG, Harlaar J. Synergy of EMG patterns in gait as an objective measure of muscle selectivity in children with spastic cerebral palsy. *Gait Posture* 2012; 35: 111–5.

Appendix 12: Publication study 2

Please see next page.

Selective voluntary motor control measures of the lower extremity in children with upper motor neuron lesions: a systematic review

JULIA BALZER¹ | MARIETTA L VAN DER LINDEN² | THOMAS H MERCER² | HUBERTUS J A VAN HEDEL¹ 

¹ Pediatric Rehabilitation Research Group, Rehabilitation Center Affoltern am Albis, University Children's Hospital Zurich, Affoltern am Albis, Switzerland. ² Centre for Health, Activity and Rehabilitation Research, Queen Margaret University, Edinburgh, UK.

Correspondence to Julia Balzer at Pediatric Rehabilitation Research Group, Rehabilitation Center Affoltern am Albis, University Children's Hospital Zurich, Mühlebergstrasse 104, CH-8910 Affoltern am Albis, Switzerland. E-mail: julia.balzer@kispi.uzh.ch

PUBLICATION DATA

Accepted for publication 18th January 2017.

Published online 8th March 2017.

ABBREVIATIONS

COSMIN	COnsensus-based Standards for the selection of health Measurement INstruments
SCALE	Selective Control Assessment of the Lower Extremity
SMC	Selective Motor Control test
SVMC	Selective voluntary motor control
UMN	Upper motor neuron

AIM Recovery and trainability of impaired selective voluntary motor control (SVMC) of the lower extremity in children with upper motor neuron lesions has received little attention. To facilitate an evidence-based debate about this topic, this review evaluates the evidence level of the psychometric properties of SVMC measures.

METHOD MEDLINE, Embase, CINAHL, PsycINFO, Scopus, Cochrane and PEDro databases were systematically searched up to July 2016. Two independent raters scored the methodological quality in accordance to the COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) checklist. The overall level of evidence was scored according to Cochrane criteria.

RESULTS We identified 3590 studies, of which 17 were included. COSMIN scores ranged from 'poor' to 'excellent' for studies investigating measurement properties of the Selective Motor Control test, modified Trost test, Gillette's Selective Motor Control test, Selective Control Assessment of the Lower Extremity (SCALE), kinematic measures, electromyography, and torque steadiness. Studies assessing the SCALE scored highest on COSMIN items. Evidence levels for SCALE's validity and reliability properties were moderate, while for the other SVMC measures these ranged from unknown to moderate. Responsiveness was not assessed.

INTERPRETATION Further psychometric studies of SVMC measures are needed to provide a scientific contribution to the ongoing debate of SVMC trainability.

Understanding and therapeutically guiding motor development and motor control is a complex and challenging topic for professionals and caregivers within the field of neuropaediatric rehabilitation.^{1–3} Motor control is a primary determinant for physiological or 'normal' movement.⁴ When measuring motor control, the following question must be considered: 'how are motor units selected, activated and deactivated?'⁵ In children with upper motor neuron (UMN) lesions, the muscle control mechanism of selected activation and deactivation is often disturbed and causes non-physiological movements (patterns).^{3,6} Clinically, this is known as a loss of selective voluntary motor control (SVMC).^{6–8} Selective motor control is defined as the ability 'to isolate the activation of muscles in a selected pattern in response to demands of a voluntary movement or posture'.⁸ The term 'voluntary' within SVMC emphasizes the deliberate performance of selected muscle activation during functional tasks.⁷

Pathophysiologically, a loss of SVMC is related to impaired descending corticospinal input, which results in disturbed control of spinal networks.^{6,8} Complex muscle activation patterns of agonist, synergist, and antagonist are

disturbed.^{8,9} This allows the appearance of flexor and/or extensor mass movement patterns (i.e. synergies), which often limits the patient's control of movement.⁷ For patients with a complete loss of SVMC, strength and functional training is only possible within synergies, potentially accompanied by co-movements and mirror movements of other muscle groups. This could limit the patient's ability to increase or maintain strength of a specific weak muscle (group).^{6–9} Therefore, in the long term, impaired SVMC can result in a vicious circle of limited motor performance during daily life activities, secondary deformities, and pain.^{6,7,10} Studies investigating the impact of different motor impairments in children with cerebral palsy (CP) have shown that a loss of SVMC limits motor performance more than other routinely measured impairments such as spasticity and/or contractures.^{10–13} Recently, selective motor control has been listed in the International Classification of Functioning, Disability and Health: Children & Youth Version (ICF) core sets for children and youth with CP (b7600),¹⁴ underscoring its clinical importance within this patient group. Furthermore, for many ambulatory children with UMN lesions and their caregivers, learning to

walk/move 'normally' (i.e. within a physiological pattern or without synergistic mass patterns) is a commonly mentioned rehabilitation goal.^{15,16} Although achieving a 'normalized' walking pattern training of multiple body functions (e.g. balance, strength) is necessary, SVMC plays a major role in performing qualitative good walking movements.¹³

Despite the social and pathophysiological importance of SVMC, evidence about its trainability in children with CP is relatively limited.^{17–20} One possible reason for the small number of intervention studies to improve SVMC might be related to the challenges of measuring motor control.⁷ In contrast to measuring motor function, which is readily done in numbers by asking 'how fast' or 'how often' a movement is performed, measuring motor control is more difficult because it looks at 'how' the movement is controlled and executed.^{1,2,7} Nevertheless, for being able to assess rehabilitation-induced (medical/therapeutic) changes of SVMC, the availability of valid, reliable, and responsive SVMC outcome measures are fundamental.

There is currently no systematic review of the psychometric properties of SVMC measures for the lower extremity for children with UMN lesions; the purpose of this study was to address this gap. By focusing on the lower extremity, we aimed to extend the observations from a recent systematic review,²¹ which investigated psychometric properties of tests scoring either compensatory or physiological movements of the upper extremity in children with UMN lesions. Providing an overview of SVMC measures for the lower extremity and the evaluation of the level of evidence of the psychometric properties of SVMC measures for children with UMN lesion will be useful for clinicians and researchers planning future studies on the trainability of SVMC.

METHOD

Search strategy

To identify studies assessing the psychometric properties of outcome measures of SVMC in children with an UMN lesion the following databases were searched without any time limit until July 2016: MEDLINE, Embase, CINAHL, PsycINFO, Scopus, Cochrane, and PEDro. The search strategy included keywords and synonyms for SVMC, as well as names of tools previously used to measure SVMC, the population of interest, and a validated search filter for finding studies on measurement properties.²² An example of the search strategy used for MEDLINE is included in Appendix S1 (online supporting information). In addition, to identify additional studies, we hand-searched the reference lists of the articles included in the review.

Study selection

We used a previously developed proprietary database (Microsoft Access 2010; Microsoft Corp., Redmond, WA, USA) to enter the data systematically and score the methodological quality of the studies.²³

What this paper adds

- Very few studies evaluate validity and reliability of selective voluntary motor control measures.
- Responsiveness studies of selective voluntary motor control measures are lacking.
- We can currently recommend the Selective Control Assessment of the Lower Extremity measure/tool.
- The evidence level of objective interval-scaled selective voluntary motor control measures is limited.

Inclusion and exclusion criteria were defined in advance. In accordance to the definition of SVMC, stated above, only papers dealing with selective movement of one joint of the lower extremity or with a primary selective (not synergistic) voluntary multijoint movement were included. For example, papers dealing with the ankle dorsiflexion during initial contact or investigating pathological synergy patterns during walking (i.e. activation of the rectus femoris and semitendinosus muscles during swing phase) were included, whereas papers measuring SVMC over the whole gait cycle or during gross motor coordination tasks were excluded. Considering that SVMC comprehends how accurately and smoothly someone can isolate the selection of a particular muscle group, papers describing the measurement of submaximal torque steadiness were included (i.e. ICF body function level b7300 power of isolated muscle activation). However, studies on maximal voluntary contraction were excluded, as patients with impaired SVMC tend to produce maximal force by using mass synergy patterns.^{24,25} Furthermore, neuroimaging measures, testing structural and metabolic intactness of the involved underlying neurophysiological structures, or networks involved in SVMC (i.e. ICF body structure levels 1100, corticospinal tract, primary cortex) were excluded. Only papers dealing with children and young people (3–21y) with UMN lesions were included. This age range was chosen for neurophysiological reasons (e.g. maturation of the corticospinal tract) and practical reasons (e.g. compliance/understanding). Studies with the explicit aim of assessing one or more psychometric properties were included, as well as cohort studies indirectly investigating the psychometric characteristics of an outcome measure by, for instance, looking at the difference between neurologically intact children and those with UMN lesions. All other forms of indirect evidence (i.e. intervention studies) were excluded. Only articles published in English and German were included for review.

Two reviewers (JB and MLvdL) independently screened all titles and abstracts of the papers. In cases of doubt, the full-text article was consulted to decide whether or not the study met the inclusion criteria. A third reviewer was available if no consensus could be achieved.

Quality evaluation

Evaluation of the methodological quality of the included papers was carried out independently by JB and MLvdL by using the four-point rating scale ('excellent', 'good', 'fair', 'poor', or 'not applicable') of the COnsensus-based Standards for the selection of health Measurement Instruments

(COSMIN) checklist.²⁶ The COSMIN checklist consists of three domains, namely validity ('The degree to which an [Health-Related Patient-Reported Outcomes (HR-PRO)] instrument measures the construct(s) it purports to measure'), reliability ('The extent to which scores for patients who have not changed are the same for repeated measurement under several conditions'), and responsiveness ('The ability of an [HR-PRO] instrument to detect change over time in the construct to be measured').²⁶ Each domain contains one or more measurement properties. The reviewer selects the measurement properties (COSMIN boxes) evaluated in the study and scores the specific item lists via the aforementioned ordinal scoring system. The lowest score of all items of the chosen COSMIN box determines the overall methodological quality of the paper. In line with previous COSMIN reviews in the field of neuropaediatrics, we adopted the overall COSMIN score by omitting the item regarding sample size.^{21,23}

To ensure that both raters scored the papers in accordance to the guidelines and to allow other raters to arrive at the same conclusion, the following procedures were established prior to the independent COSMIN rating: raters were familiarized with the COSMIN manual and terminology, and discussed the scoring of two papers and established additional rating rules (Appendix S2, online supporting information). Although the COSMIN manual provides general rules for all boxes and items, for some items the COSMIN rating is still open to subjective interpretation, for example 'time interval appropriate'. It is for this reason that COSMIN itself recommends specification of additional rules for individual reviews.²⁷ If the two reviewers could not agree on a score, a third reviewer was made available.

For the assessment of quality of the measurement properties, the updated criteria suggested by Terwee et al.²⁸ were applied (Appendix S3, online supporting information). The overall level of evidence for each SVMC measure and each measurement property was evaluated according to the Cochrane Back Review Group Criteria:²⁹ 'strong', 'moderate', 'limited', 'conflicting', 'unknown' (Appendix S4, online supporting information). This overall score was given in relation to the methodological quality of the study and the results of the measurement properties.²⁶ Again, criteria for sample size were adapted as follows: sample sizes greater than 100 participants of the combined studies were rated as 'strong' (+++ or ---); studies with a sample size between 50 and 99 were rated as 'moderate' (++ or --); studies with a sample size between 25 and 59 as 'limited' (+ or -); and sample size less than 25 as 'unknown' (?).³⁰

RESULTS

Description of the included studies

The systematic search resulted in 3590 references being identified. Based on the titles and abstracts, 33 papers were included for full-text reading. After applying the inclusion and exclusion criteria, 17 papers were retained for review (Figure S1, online supporting information).

These 17 papers described the measurement properties of four clinical, ordinal-scaled, assessment tools (Selective Motor Control test [SMC]; modified Trost test [mTrost]; Gillette's SMC test; Selective Control Assessment of the Lower Extremity [SCALE]) and three laboratory based interval-scaled measurement tools (kinematic measures, electromyography [EMG], and torque steadiness). The majority of studies tested SVMC of the ankle or the knee joint.

The following psychometric properties were evaluated: hypotheses testing/construct validity was assessed in 17 studies, reliability in six (interrater, $n=5$; test-retest, $n=3$; intrarater, $n=1$), and both content and criterion validity in one study. Responsiveness was not evaluated in any study. Most studies tested the SCALE ($n=9$), followed by studies evaluating torque steadiness measures ($n=2$), kinematic measures ($n=4$), and EMG of selected lower limb muscles ($n=1$). The age of the participants in the studies included for final review ranged from 2 to 21 years, with the exception of one study,³¹ where the oldest participant was 28 years of age. Although this age range was slightly wider than the one set by the inclusion criteria (3–21y), discussing this issue ended in the common decision for inclusion. As the mean age ranged from 9 years 3 months to 16 years, the youngest and the oldest participant were seen as outliers. Sample size varied from eight to 51 participants. All studies included children with a diagnosis of CP. In two cohort studies data of children with CP who had undergone a selective dorsal rhizotomy were compared with those of a control group children with CP who had not.^{32,33} Selective dorsal rhizotomy is a neurosurgical procedure, which aims to minimize the limiting influence of spasticity on motor control in children with spastic CP. The comparison of SVMC between children with and without rhizotomy was therefore considered an assessment of the validity of the SVMC tool. Four other cohort studies investigated the construct validity of the SVMC instrument,^{31,34–36} by comparing patients with CP with participants who were neurologically intact. General characteristics and clinical utility for each SVMC measure is summarized in Table SI (online supporting information). The methodological quality per measurement property, as well as the overall evidence criteria, can be seen in Table SII (online supporting information).

Hypotheses testing

Of the 17 papers that evaluated construct validity ('hypotheses testing'), 10 papers included clinical assessment tools and seven papers laboratory-based measurement tools (Tables SI and SII).

Nine of the 10 papers regarding clinical assessment tools evaluated construct validity of the SCALE. The modified COSMIN scores of three of these eight SCALE papers were 'good',^{37–39} four were rated as 'fair',^{40–42} one as 'excellent',⁴³ and one as 'poor'.⁴⁴ Quality of construct validity was evaluated in accordance to Terwee et al.²⁸ as 'positive' (+) in eight papers,^{37–42,45} and as mixed 'positive/negative' (+/–) in one study.⁴⁴ Overall, there was 'moderate positive' (++)

evidence for construct validity of the SCALE in terms of (1) its correlation with the Gross Motor Function Classification System,^{37,43} and (2) its proximal–distal concordance (SVMC is more often and/or more severely impaired in distal body parts).^{38,43} A 'limited positive (+)' evidence level was given for its validity testing with the Berg Balance Scale,⁴⁵ and for predicting knee flexion during initial contact during stance phase of gait.⁴² Three studies investigated relationships between total limb SCALE scores and knee flexion during swing phase. Two studies found significant correlations and one did not.^{39–41} Therefore, their level of evidence was rated as 'conflicting (–)'. In relation to the poor quality of the Kusumoto et al. study,⁴⁴ which investigated the relationship between the SCALE and knee extensor strength, its level of evidence was rated as 'unknown (?)'. Therefore, this study did not contribute to the overall evidence level of the SCALE.

In the other three SVMC clinical assessment tools construct validity was only evaluated for the Gillette's SMC test. The study of Manikowska et al. compared, using EMG, Gillette's knee flexion SMC scores in patients with CP versus participants who were neurologically intact.³¹ This study received a 'poor' modified COSMIN score, but results were rated as 'positive (+)' in accordance to Terwee et al.²⁸ The level of evidence was evaluated as 'unknown (?)'.³⁰

Seven papers investigated construct validity of SVMC using laboratory-based measurement tools, including lower limb kinematics,^{32,34,39,40} EMG,³³ and torque steadiness.^{35,36} Two kinematic papers were already evaluated in relation to the SCALE's construct validity (see previous paragraph),^{39,40} but we decided to list them under this section as well. As there is no criterion standard measure for quantifying SVMC, and the papers are cohort studies investigating the correlation between the SCALE and a kinematic measure for SVMC, they could be regarded not only as studies investigating the construct validity of the SCALE, but also of the kinematic measures, depending on which measure is regarded as more 'established'. As the quality and evidence rating of the two studies is the same as presented above,^{39,40} the results will not be repeated here. Their COSMIN methodological quality was rated as 'fair' and quality of construct validity was rated as 'positive (+)' in three studies,^{34–36} and as mixed 'positive/negative (+)/(–)' in two.^{32,33} Because the sample size was too small ($n < 25$) in three studies, the evidence level was only scored for one kinematic and one EMG study.^{33,34} Because the results of these studies were ambiguous in supporting the construct validity of the SVMC measurement method, a 'conflicting (–)' evidence rating was assigned.

Overall, the methodological quality of the majority of the abovementioned studies was reduced owing to the absence of a priori formulated hypotheses, thereby limiting their COSMIN,²⁷ as well as validity, quality scoring.²⁹

Content and criterion validity

Content and criterion validity were only assessed for the SCALE (Table SII).^{37,43} COSMIN rating of content

validity was considered 'poor'.²⁷ Although 14 experts were involved on item agreement for statements about content, administration, and grading of the SCALE, the paper lacked a description of whether all items were relevant for the construct or for the population of interest.³⁷ The quality rating of the results was scored as 'indeterminate (?)' and the evidence as 'unknown (?)'.²⁷

The method applied to establish criterion validity of the SCALE was rated as 'excellent'.⁴³ As the Fugl–Meyer Assessment (items III and IV) measures a similar construct as the SCALE and correlation exceeded 0.70, the SCALE criterion validity results were rated as 'positive (+)'. Therefore, a 'limited positive (+)' evidence level was given for criterion validity of the SCALE.

Reliability

Reliability was investigated in three (SMC, mTrost, SCALE)^{37,43,46,47} of the four clinical assessment tools and in two (kinematic, torque)^{32,35} of five laboratory-based SVMC tools. The SMC test–retest reliability was tested in two studies.^{46,47} The modified COSMIN rating for SMC interrater reliability ranged from 'fair' to 'good'.^{27,46,47} The methodological quality of interrater reliability of the mTrost test was also rated 'fair'.⁴⁷ Interrater reliability of the SCALE was tested by two studies and scored as 'excellent' and 'good'.^{37,43} The SCALE's intratester reliability was further investigated and received a 'good' modified COSMIN score.⁴³ The methodological quality of test–retest reliability for the kinematic and torque steadiness measure was evaluated as 'good'.^{32,35} Overall, studies assessing interrater reliability were rated lowest for COSMIN items describing the statistical procedures (i.e. description of weighted scheme intraclass correlation coefficient, kappa). The items regarding the stability of participants between the two or more assessments and the description of test conditions were the most limiting items for the four studies on test–retest and intratester reliability.

Applying the quality criteria for measurement properties revealed 'positive (+)' results for four reliability studies;^{32,37,43,46} mixed 'positive/negative (+)/(–)' results in three studies;^{35,46,47} and 'negative (–)' results for interrater reliability of the SMC test.²⁹

When evaluating the overall evidence level using Cochrane guidelines,³⁰ we found 'moderate positive results (++)' for the inter- and intratester reliability of the SCALE.^{37,43} 'Moderate negative results (–)' were evident for the interrater reliability of the SMC,^{46,47} and 'limited negative results (–)' for the mTrost.⁴⁷ Owing to the low sample size ($n < 25$) the evidence level of the test–retest reliability of the kinematic and torque steadiness measurement studies was scored as 'unknown (?)'.^{32,35}

DISCUSSION

This review revealed a limited number of psychometric studies investigating SVMC measures in children with UMN lesions. The overall evidence was further limited as 10 out of 17 studies were cohort studies with a limited methodological quality (i.e. 'poor' or 'fair'), with the

exception of one study,³⁹ which was scored 'good' according to modified COSMIN rating guidelines. No study investigated responsiveness, which we consider a crucially important measurement property, especially in the context of the ongoing debate about trainability of SVMC. Although we explicitly searched for the whole population of paediatric UMN lesions, only psychometric studies including children with CP were found. Therefore, the results of this review are limited to children with CP and cannot be directly transferred to other children with impaired SVMC (i.e. acquired brain injuries). The chosen age range (3–21y) for this review, might have been wide when considering developmental issues known to influence SVMC (e.g. maturation of central nervous system function), as well as the importance of the participants' cognitive understanding and motivation for the SVMC measurement procedure/testing. Nevertheless, this age range was chosen owing to the overall limited number of studies available for review. Future studies regarding SVMC measures may choose to investigate psychometric properties in separate age groups (i.e. pre-central nervous system maturation 2–7y, and post-central nervous system maturation <8y).

The SCALE was the most often investigated assessment tool, in terms of number of studies conducted and in the number of measurement properties investigated. The SMC and mTrost were only rated on reliability, thus lacking evidence on their validity. The Gillette's SMC test was only investigated on its validity, lacking evidence about its reliability. In terms of psychometric quality, the SCALE had the highest level of evidence with a moderate positive level of evidence concerning its interrater reliability and its construct validity, and an unknown and limited level of evidence of content and criterion validity respectively. Another advantage of the SCALE over the other assessment tools (SMC, mTrost) lies in its evaluation of five lower extremity joints rather than one or three joints. In addition, clinical utility (Table S1) of the SCALE, as well as of the other SMC assessment tools, was scored high as required time, costs, and resources are low. However, in terms of limitations, the SCALE's application is limited to children with spastic CP. Furthermore, as its ordinal scoring system relies on the impression of the rater (i.e. therapist, consultant), it is a subjective measurement. Finally, the SCALE's three-level ordinal scoring system (normal, impaired, and unable) may lack sensitivity to detect certain therapy-induced changes of SVMC. These limitations (spastic CP, ordinal scoring system) also apply for the other three clinical tools.

The construct validity of the kinematic, EMG, and torque steadiness was assessed, but none of the papers evaluating these measurement techniques explicitly mentioned that the assessment of validity was an *a priori* objective. Because of this, the formulation of hypotheses was often absent thus diminishing their modified COSMIN score to 'fair'. Only two laboratory-based SVMC measures (kinematic, torque steadiness)^{32,35} were assessed regarding their

test-retest reliability. In terms of psychometric quality, as well as clinical utility (Table S1), none of the identified laboratory-based measures seem to offer great advantage over the other. The equipment required to record the outcome measures was often customized, making it difficult for other groups (researchers or clinicians) to apply and confirm or extend findings of studies exploring the laboratory-based measures using EMG, kinematics, or torque measurements. Furthermore, the measurement procedures appear time-consuming and complex in comparison with more routinely applied clinical assessments. Personnel also required extensive training in the application and analysis of these measures (Table S1).

In summary, the results from this systematic review show the limited level of evidence regarding the psychometric properties (reliability and validity) and absence of evidence regarding responsiveness of currently available SVMC measures of the lower extremity in children with UMN lesions.

Methodological considerations

Low interrater agreement when rating the quality of the evidence in systematic reviews (e.g. rating risk of bias in Cochrane type reviews) can be an important methodological issue, which should be considered when conducting a systematic review.⁴⁸ Herein, agreement between the raters for all COSMIN items was high. Only five out of 246 items needed further discussion, and none required the rating of a third reviewer. This high agreement was likely the result of the specific rating rules which we established as recommended by the COSMIN group. For example, when scoring the reliability items 4 to 7 for the SCALE, we decided in advance to score the use of video for the evaluating of the inter- and intrarater reliability as appropriate, as this allows a discrete evaluation of the scoring system by maintaining the stability of test conditions and patient status, as well as saving on time and resources. In contrast, a video approach was not considered to be appropriate for determining test-retest reliability when the stability of the patient is evaluated.

In line with other neuropaediatric COSMIN reviews,^{22,24,49} we modified the rating of the sample size item (modified COSMIN score) as the sample size is often limited in clinical neuropaediatric studies and not comparable with large-scale epidemiological healthcare studies using patient-reported outcome measures for which the COSMIN guidelines were initially evaluated. This modified scoring improved the overall rating of all studies with the exception of the construct and content validity score of the studies from Zwaan et al.³³ and Fowler et al.³⁷ Although we used this modified score, we recommend that future psychometric studies should include a sufficiently large sample size (>30).

Other reasons for scoring poor were the lack of 'a priori' formulated hypotheses (box 'hypothesis testing'), and for one study the lack of evaluating each item separately for its content validity (box 'content validity').³⁷ While we consider

it important that each single question should be evaluated separately for its content validity in a healthcare questionnaire (which the COSMIN was originally developed for), it could be questioned whether the same rating rules are needed for an assessment tool like the SCALE that consists of a similar procedure repeated for different joints.

Future research directions

The results of this review show that the SCALE is the most frequently investigated assessment method in the population of children with CP and also has the highest quality rating. Its responsiveness to change has not been assessed, but it may be expected that owing to its ordinal scoring system its sensitivity to measure changes of SVMC is limited. To improve its sensitivity and, simultaneously, to benefit from its child-friendly procedure, combining the SCALE with another, more sensitive measure appears to be promising. This idea has also been proposed in previous studies.^{33,50} While Zwaan et al. found no convincing evidence for detecting extensor and flexor synergies during gait using EMG in children with CP,³³ they reported a significant cross-correlation between extensor synergy activity measured using EMG and the mTrost test. They concluded that 'EMG measures still may be useful for selective motor control measurement because it measures selectivity at the level of the specific muscles involved, provided the appropriate task is used'.³³ As walking requires selective, as well as synergistic movements, we considered it not to be an appropriate task for the assessment of SVMC. The tasks embedded in the SCALE (isolated single-joint movements) were developed in accordance with the definition of SVMC.^{8,37} Combining SCALE's ratings for single-joint movements with EMG recordings would further allow for directly measuring voluntary activation of a muscle, even in patients with low muscle strength (manual muscle test grade of 1), where no real joint movement occurs.

CONCLUSION

This systematic review revealed a limited number of psychometric studies evaluating the validity and reliability of SVMC measures in children with UMN lesions, and no

studies evaluating responsiveness. Currently, the SCALE appeared to have the highest level of evidence regarding its reliability and construct validity compared with other clinical and laboratory-based measures of SVMC. However, only by means of reliable, validated, and responsive SVMC tools used in carefully designed intervention studies, will it be possible to provide a scientifically rigorous contribution to the ongoing debate with regard to the possibility to improve SVMC of the lower extremity in children with UMN lesions.

ACKNOWLEDGEMENTS

This research project was funded by the Swiss National Science Foundation (grant number 32003B_156646), 'PhysioSwiss' (Switzerland), 'Physiotherapy Science Foundation' (Switzerland), and 'Mexi-Foundation' (Switzerland). We further acknowledge the Zurich Center for Neuroscience. None of the funders were involved in the study design, data collection, analysis, and manuscript preparation and publication decisions. All ideas and decisions in relation to this study were made independently by the authors. The authors have stated that they had no interests which might be perceived as posing a conflict or bias.

SUPPORTING INFORMATION

The following additional material may be found online:

Appendix S1: Search construct: psychometric studies of selective voluntary motor control measures for the lower extremity in children with upper motor neuron lesions.

Appendix S2: Specified COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) rating rules.

Appendix S3: Quality criteria for measurement properties.

Appendix S4: Levels of evidence for the overall quality of the measurement properties (based on the Cochrane Back Review Group 2003).

Figure S1: Selection process systematic review. SVMC, selective voluntary motor control.

Table S1: General characteristics, psychometric properties, and clinical utility of selective voluntary motor control measures.

Table SII: Summary study details, COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN)-, quality-, and evidence-rating.

REFERENCES

- Boyce WF, Gowland C, Hardy S, et al. Development of a quality-of-movement measure for children with cerebral palsy. *Phys Ther* 1991; 71: 820–8.
- Rosenbaum PL, Russell DJ, Cadman DT, Gowland C, Jarvis S, Hardy S. Issues in measuring change in motor function in children with cerebral palsy: a special communication. *Phys Ther* 1990; 70: 125–31.
- Campbell SK. Assessment of the child with central nervous system dysfunction. In: Rothstein JM, editor. *Measurement in Physical Therapy*. New York, NY: Churchill Livingstone, 1985: 207–28.
- Rothwell JC. *Control of Human Voluntary Movement*. London: Croom Helm, 1987.
- Rose J, McGill KC. Neuromuscular activation and motor-unit firing characteristics in cerebral palsy. *Dev Med Child Neurol* 2005; 47: 329–36.
- Cahill-Rowley K, Rose J. Etiology of impaired selective motor control: emerging evidence and its implications for research and treatment in cerebral palsy. *Dev Med Child Neurol* 2014; 56: 522–8.
- Fowler EG. Concepts in spasticity and selective motor control in children with spastic cerebral palsy. *Technol Disabil* 2010; 22: 207–14.
- Sanger TD, Chen D, Delgado MR, Gaebler-Spira D, Hallett M, Mink JW. Definition and classification of negative motor signs in childhood. *Pediatrics* 2006; 118: 2159–67.
- Rose J. Selective motor control in spastic cerebral palsy. *Dev Med Child Neurol* 2009; 51: 578–9.
- Vos RC, Becher JG, Voorman JM, et al. Longitudinal association between gross motor capacity and neuromusculoskeletal function in children and youth with cerebral palsy. *Arch Phys Med Rehabil* 2016; 97: 1329–33.
- Østensjø S, Carlberg EB, Vøllestad NK. Motor impairments in young children with cerebral palsy: relationship

- to gross motor function and everyday activities. *Dev Med Child Neurol* 2004; 9: 580–9.
12. Voorman JM, Dallmeijer J, Knol DL, Lankhorst GJ, Becher JG. Prospective longitudinal study of gross motor function in children with cerebral palsy. *Arch Phys Med Rehabil* 2007; 88: 871–6.
 13. Desloovere K, Molenaers G, Feys H, Huenaerts C, Callewaert B, Van de Walle P. Do dynamic and static clinical measurements correlate with gait analysis parameters in children with cerebral palsy? *Gait Posture* 2006; 24: 302–13.
 14. Schiari V, Masse LC. Relevant areas of functioning in children with cerebral palsy based on the International Classification of Functioning, Disability and Health Coding System: a clinical perspective. *J Child Neurol* 2014; 30: 216–22.
 15. Wu YN, Hwang M, Ren Y, Gaebler-Spira D, Zhang LQ. Combined passive stretching and active movement rehabilitation of lower-limb impairments in children with cerebral palsy using a portable robot. *Neurorehabil Neural Repair* 2011; 25: 378–85.
 16. Beveridge B, Feltracco D, Struyf J, et al. “You gotta try it all”: parents’ experiences with robotic gait training for their children with cerebral palsy. *Phys Occup Ther Pediatr* 2015; 35: 327–41.
 17. Rios DC, Gilbertson T, McCoy SW. NeuroGame Therapy to improve wrist control in children with cerebral palsy: a case series. *Dev Neurorehabil* 2013; 16: 398–409.
 18. Sukal-Moulton T, Clancy T, Zhang LQ, Gaebler-Spira D. Clinical application of a robotic ankle training program for cerebral palsy compared to the research laboratory application: does it translate to practice? *Arch Phys Med Rehabil* 2014; 95: 1433–40.
 19. Burdea GC, Cioi D, Kale A, Janes WE, Ross SA, Engsberg JR. Robotics and gaming to improve ankle strength, motor control, and function in children with cerebral palsy: a case study series. *IEEE Trans Neural Syst Rehabil Eng* 2013; 21: 165–73.
 20. Van Doornik J, Kukke S, McGill K, Rose J, Sherman-Levine S, Sanger TD. Oral baclofen increases maximal voluntary neuromuscular activation of ankle plantar flexors in children with spasticity due to cerebral palsy. *J Child Neurol* 2008; 23: 635–9.
 21. Gerber CN, Labruyere R, van Hedel HJ. Reliability and responsiveness of upper limb motor assessments for children with central neuromotor disorders: a systematic review. *Neurorehabil Neural Repair* 2016; 30: 19–39.
 22. Terwee CB, Jansma EP, Riphagen II, De Vet HCW. Development of a methodological PubMed search filter for finding studies on measurement properties of measurement instruments. *Qual Life Res* 2009; 18: 1115–23.
 23. Ammann-Reiffer C, Bastiaenen CHG, de Bie R, van Hedel HJ. Measurement properties of gait-related outcomes in youth with neuromuscular diagnoses: a systematic review. *Phys Ther* 2014; 94: 1067–82.
 24. Damiano DL, Dodd K, Taylor NF. Should we be testing and training muscle strength in cerebral palsy? *Dev Med Child Neurol* 2002; 44: 68–72.
 25. Tedroff K, Knutson LM, Soderberg GL. Synergistic muscle activation during maximum voluntary contractions in children with and without spastic cerebral palsy. *Dev Med Child Neurol* 2006; 48: 789–96.
 26. Mokkink LB, Terwee CB. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. *Qual Life Res* 2010; 19: 539–49.
 27. Mokkink LB, Terwee CB, Gibbons E, et al. Inter-rater agreement and reliability of the COSMIN (Consensus-based Standards for the selection of health status Measurement Instruments) checklist. *BMC Med Res Methodol* 2010; 10: 82.
 28. Terwee CB, Bot SD, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol* 2007; 60: 34–42.
 29. Van Tulder M, Furlan A, Bombardier C, Bouter L; Editorial Board of the Cochrane Collaboration Back Review Group. Updated method guidelines for systematic reviews in the Cochrane collaboration back review group. *Spine* 2003; 28: 1290–9.
 30. Dobson F, Hinman RS, Hall M, Terwee CB, Roos EM, Bennell KL. Measurement properties of performance-based measures to assess physical function in hip and knee osteoarthritis: a systematic review. *Osteoarthritis Cartilage* 2012; 20: 1548–62.
 31. Manikowska F, Chen B, Jozwiak M, Lebidowska MK. Assessment of selective motor control in clinical Gillette’s test using electromyography. *Eur J Phys Rehabil Med* 2016; 52: 176–85.
 32. Engsberg JR, Ross SA, Park TS. Quantifying active ankle range of motion in cerebral palsy following selective dorsal rhizotomy. *J Appl Biomech* 2004; 20: 103–11.
 33. Zwaan E, Becher JG, Harlaar J. Synergy of EMG patterns in gait as an objective measure of muscle selectivity in children with spastic cerebral palsy. *Gait Posture* 2012; 35: 111–5.
 34. Engsberg JR, Van Emmerik RE, Ross SA, Collins DR. Use of relative phase as a measure of motor control at the ankle in persons with cerebral palsy: a preliminary study. *J Appl Biomech* 2008; 24: 382–90.
 35. Bandholm T, Rose MH, Sløk R, Sonne-Holm S, Jensen BR. Ankle torque steadiness is related to muscle activation variability and coactivation in children with cerebral palsy. *Muscle Nerve* 2009; 40: 402–10.
 36. Arpin DJ, Stueberg W, Stergiou N, Kurz MJ. Motor control of the lower extremity musculature in children with cerebral palsy. *Res Dev Disabil* 2013; 34: 1134–43.
 37. Fowler EG, Staudt L, Greenberg MB, Oppenheim WL. Selective Control Assessment of the Lower Extremity (SCALE): development, validation, and interrater reliability of a clinical tool for patients with cerebral palsy. *Dev Med Child Neurol* 2009; 51: 607–14.
 38. Fowler EG, Staudt L, Greenberg MB. Lower-extremity selective voluntary motor control in patients with spastic cerebral palsy: increased distal motor impairment. *Dev Med Child Neurol* 2010; 52: 264–9.
 39. Goldberg EJ, Requejo PS, Fowler EG. Joint moment contributions to swing knee extension acceleration during gait in children with spastic diplegic cerebral palsy. *Gait Posture* 2011; 33: 66–70.
 40. Fowler EG, Goldberg EJ. The effect of lower extremity selective voluntary motor control on interjoint coordination during gait in children with spastic diplegic cerebral palsy. *Gait Posture* 2009; 29: 102–7.
 41. Rha DW, Cahill-Rowley K, Young J, Torburn L, Stephenson K, Rose J. Biomechanical and clinical correlates of swing-phase knee flexion in individuals with spastic cerebral palsy who walk with flexed-knee gait. *Arch Phys Med Rehabil* 2015; 96: 511–7.
 42. Rha DW, Cahill-Rowley K, Young J, Torburn L, Stephenson K, Rose J. Biomechanical and clinical correlates of stance-phase knee flexion in persons with spastic cerebral palsy. *PMR* 2016; 8: 11–8.
 43. Balzer J, Marsico P, Mitteregger E, van der Linden ML, Mercer TH, van Hedel HJ. Construct validity and reliability of the Selective Control Assessment of the Lower Extremity in children with cerebral palsy. *Dev Med Child Neurol* 2016; 58: 167–72.
 44. Kusumoto Y, Takaki K, Matsuda T, Nitta O. Relation of selective voluntary motor control of the lower extremity and extensor strength of the knee joint in children with spastic diplegia. *J Phys Ther Sci* 2016; 28: 1868–71.
 45. Lim H. Correlation between the selective control assessment of lower extremity and pediatric balance scale scores in children with spastic cerebral palsy. *J Phys Ther Sci* 2015; 27: 3645–9.
 46. Lëwing K, Carlberg EB. Reliability of the selective motor control scale in children with cerebral palsy. *Adv Physiother* 2009; 11: 58–63.
 47. Smits D-W, van Groenestijn AC, Ketelaar M, Scholtes V, Becher JG, Gorter JW. Selective motor control of the lower extremities in children with cerebral palsy: inter-rater reliability of two tests. *Dev Neurorehabil* 2010; 13: 258–65.
 48. Jørgensen L, Paludan-Møller AS, Laursen DR, et al. Evaluation of the Cochrane tool for assessing risk of bias in randomized clinical trials: overview of published comments and analysis of user practice in Cochrane and non-Cochrane reviews. *Syst Rev* 2016; 5: 80.
 49. Lennon N, Thorpe D, Balemans AC. The clinimetric properties of aerobic and anaerobic fitness measures in adults with cerebral palsy: a systematic review of the literature. *Res Dev Disabil* 2015; 45–46: 316–32.
 50. Dobson F. Assessing selective motor control in children with cerebral palsy. *Dev Med Child Neurol* 2010; 52: 409–10.